

ORIGINAL

414 Rec'd PCT/PTO 2 2 JUN 2000

FORM PTO-1390
(REV 11-98)

U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

ATTORNEY'S DOCKET NUMBER

**TRANSMITTAL LETTER TO THE UNITED STATES
DESIGNATED/ELECTED OFFICE (DO/EO/US)
CONCERNING A FILING UNDER 35 U.S.C. 371**

LEA 32 805

U.S. APPLICATION NO. (if known see 37 CFR 1.5)

09/582246

INTERNATIONAL APPLICATION NO.

PCT/EP98/08216

INTERNATIONAL FILING DATE

22 December 1998 (22.12.98)

PRIORITY DATE CLAIMED

24 December 1997 (24.12.97)

TITLE OF INVENTION

REGULATORY DNA SEQUENCES OF THE HUMAN CATALYTIC RELOMERASE SUB-UNIT GENE,
DIAGNOSTIC AND THERAPEUTIC USE THEREOF

APPLICANT(S) FOR DO/EO/US

HAGEN, Gustav; WICK, Maresa; and ZUBOV, Dmitry

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☒ This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(h) and PCT Articles 22 and 39(i).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
 - a. ☒ is transmitted herewith (required only if not transmitted by the International Bureau).
 - b. ☐ has been transmitted by the International Bureau.
 - c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
- ☒ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
- ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
 - a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
 - b. ☐ have been transmitted by the International Bureau.
 - c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
 - d. ☒ have not been made and will not be made.
- ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
- ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
- ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern document(s) or information included:

11. ☒ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. ☒ A FIRST preliminary amendment.
☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information:
 - 1) Certification of Mailing under 37 C.F.R. 1.10;
 - 2) Transmittal of Information Disclosure Statement;
 - 3) Information Disclosure Citation (Modified Form PTO-1449);
 - 4) References cited; and
 - 5) Return Receipt Post Card.

Date of Deposit: 22 June 2000

Express Mail Label No. EF292675302US

U.S. APPLICATION NO. (if known see 31 CFR 1.55)

INTERNATIONAL APPLICATION NO.
PCT/EP98/08216ATTORNEY'S DOCKET NUMBER
LEA 32 805

09/582246

17. ☒ The following fees are submitted:**BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5))**

Neither international preliminary examination fee (37 CFR 1.482)
nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO
and International Search Report not prepared by the EPO or JPO \$970.00

International preliminary examination fee (37 CFR 1.482) not paid to
USPTO but International Search Report prepared by the EPO or JPO \$840.00

International preliminary examination fee (37 CFR 1.482) not paid to USPTO but
international search fee (37 CFR 1.445(a)(2)) paid to USPTO \$760.00

International preliminary examination fee paid to USPTO (37 CFR 1.482)
but all claims did not satisfy provisions of PCT Article 33(I)-(4) \$670.00

International preliminary examination fee paid to USPTO (37 CFR 1.482)
and all claims satisfied provisions of PCT Article 33(I)-(4) \$96.00

ENTER APPROPRIATE BASIC FEE AMOUNT =**CALCULATIONS** PTO USE ONLY

\$ 840.00

Surcharge of \$130.00 for furnishing the oath or declaration later than ☐ 20 ☐ 30
months from the earliest claimed priority date (37 CFR 1.492(c)).

\$

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE
Total claims	12 -20 =	0	X \$18.00
Independent claims	9 -3 =	6	X \$78.00
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ \$260.00

\$ 00.00

\$ 468.00

\$ 0.00

TOTAL OF ABOVE CALCULATIONS = \$ 1,308.00

Reduction of 1/2 for filing by small entity, if applicable. A Small Entity Statement
must also be filed (Note 37 CFR 1.9, 1.27, 1.28).

\$ 0.00

SUBTOTAL = \$ 1,308.00

Processing fee of \$130.00 for furnishing the English translation later than ☐ 20 ☐ 30
months from the earliest claimed priority date (37 CFR 1.492(f)).

\$

+

TOTAL NATIONAL FEE = \$ 1,308.00

Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be
accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property

\$

+

\$ 0.00

TOTAL FEES ENCLOSED = \$ 1,308.00**Amount to be:****refunded** \$**charged** \$a. ☐ A check in the amount of \$_____ to cover the above fees is enclosed.b. ☒ Please charge my Deposit Account No. 13-3372 in the amount of \$ 1,308.00 to cover the above fees.
A duplicate copy of this sheet is enclosed.c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any
overpayment to Deposit Account No. 13-3372. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR
1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO

Jeffrey M. Greenman
Vice President, Patents and Licensing
BAYER CORPORATION
400 Morgan Lane
West Haven, CT 06516
US

Jerrie L. Chiu
SIGNATURE

Jerrie L. Chiu

NAME

41,670

REGISTRATION NUMBER

09/582246

534 Rec'd PCT/PTC 22 JUN2000
PATENT

Attorney's Docket No. Le A 32 805

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: Hagen, et al.

Serial No.: National Stage Filing of PCT/EP98/08216

Filed: 22 June 2000

For: Regulatory DNA Sequences of the Human Catalytic Telomerase Sub-unit Gene, Diagnostic and Therapeutic Use Thereof

BOX PCT

Assistant Commissioner for Patents
Washington, D.C. 20231

CERTIFICATE OF MAILING UNDER 37 CFR 1.10

I hereby certify that the *attached* correspondence comprising:

- Transmittal Letter to the United States Designated/Elected Office (DO/EO/US) Concerning a Filing under 35 U.S.C. 371 [IN DUPLICATE];
- A First Preliminary Amendment;
- Combined Declaration and Power of Attorney (35 U.S.C. 371(c)(4));
- English translation of the International Application (35 U.S.C. 371(c)(2));
- Copy of the International Application as filed (35 U.S.C. 371(c)(2));
- Information Disclosure Statement under 37 C.F.R. 1.97 and 1.98 consisting of Transmittal of Information Disclosure Statement, Information Disclosure Citation (Modified Form PTO-1449), and copies of references cited therein; and
- Return Receipt Post Card.

is, on the date shown below, being deposited with the United States Postal Service, in an envelope as "Express Mail Post Office to Addressee" Mailing Label Number EF292675302US, addressed to:

Box PCT
Assistant Commissioner for Patents
Washington, D.C. 20231

22 June 2000

Date


Signature of Person Certifying: Lauren Fitzgerald

09/582246

PATENT

Atty. Docket No.: Le A 32 805

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

534 Rec'd PCT/PTC 22 JUN 2000

APPLICANTS: Hagen, *et al.*

SERIAL NO.: National Stage Filing of PCT/EP98/08216

FILING DATE: Herewith

TITLE: Regulatory DNA Sequences of the Human Catalytic Telomerase Sub-Unit Gene, Diagnostic Therapeutic Use Thereof

PRELIMINARY AMENDMENT

Box PCT
Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

This Preliminary Amendment is submitted in the above-captioned national stage application of PCT/EP98/08216 filed on even date herewith. Please amend the application as follows:

In the Claims

Please cancel claim 7.

Please amend claims 4, 6 and 8-12 as follows:

4. (Amended) Recombinant construct which contains a DNA sequence according to [one of] Claim[s] 1 [to 3].
6. (Amended) Vector which contains a recombinant construct according to Claim 4 [or 5].
8. (Amended) Recombinant host cells which harbour recombinant constructs or vectors according to [one of] Claim[s] 4 [to 6].

09/582246-106170

- 0503246-092100
9. (Amended) Process for identifying substances which affect the promoter activity, silencer activity or enhancer activity of the human catalytic telomerase subunit, comprising the following steps:
 - A. adding a candidate substance to a host cell which harbours DNA sequences according to [one of] Claim[s] 1 [to 3] which sequences are functionally linked to a reporter gene, and
 - B. measuring the effect of the substance on expression of the reporter gene.
 10. (Amended) Process for identifying factors which bind specifically to the DNA according to [one of] Claim[s] 1 [to 3], or to fragments thereof, characterized in that an expression cDNA library is screened using a DNA sequence according to [one of] Claim[s] 1 [to 3], or sub-fragments of widely differing length, as the probe.
 11. (Amended) Transgenic animals which harbour recombinant constructs or vectors according to Claim[s] 4 [to 6].
 12. (Amended) Process for detecting telomerase-associated conditions in a patient, comprising the following steps:
 - A. incubating a recombinant construct or vector according to Claim[s] 4 [to 6], which additionally contains a reporter gene, with body fluids or cell samples,
 - B. detecting the activity of the reporter gene in order to obtain a diagnostic value, and
 - C. comparing the diagnostic value with standard values for the reporter gene construct in standardized normal cells or body fluids of the same type as the test sample.

Please add the following new claim 13.

13. (New) A medicament comprising a recombinant construct or vector according to claim 4.

Remarks

By way of this Preliminary Amendment, claims 1-6 and 8-13 are pending in the application. Claims 4, 6 and 8-12 have been amended. Claim 13 has been added. These claim amendments, cancellations and additions are being made solely to remove multiple claim dependencies from the claims and to place the claims in a format appropriate for U.S. prosecution.

Applicants believe that the subject matter of the pending claims is patentable and that the instant application should accordingly be allowed. If the Examiner believes that a conversation with Applicants' attorney would be helpful in expediting prosecution of this application, the Examiner is invited to call the undersigned attorney at (203) 812-3964.

Respectfully submitted,

Dated:

June 22, 2000

Bayer Corporation
400 Morgan Lane
West Haven, CT 06516
(Tel) (203) 812-3964
(Fax) (203) 812-5492
e-mail: jerrie.chiu.b@bayer.com

Jerrie L. Chiu
Jerrie L. Chiu
Attorney for Applicants
Reg. No. 41,670

Regulatory DNA sequences of the gene for the human catalytic telomerase subunit, and their diagnostic and therapeutic use

Structure and function of the chromosome ends

5

The genetic material of eukaryotic cells is distributed on linear chromosomes. The ends of hereditary units are termed telomeres, derived from the Greek words *telos* (end) and *meros* (part, segment). Most telomeres consist of repeats of short sequences which are mainly composed of thymine and guanine (Zakian, 1995). In all the vertebrates which have so far been investigated, the telomeres consist of the sequence TTAGGG (Meyne *et al.*, 1989).

10

15

The telomeres have a variety of important functions. They prevent the fusion of chromosomes (McClintock, 1941) and thus the formation of dicentric hereditary units. Such chromosomes having two centromeres can lead to the development of cancer due to loss of heterozygosis or duplication, or loss of genes.

20

In addition, telomeres serve the purpose of distinguishing intact hereditary units from damaged hereditary units. Thus, yeast cells ceased their cell division when they contained a chromosome without a telomere (Sandell and Zakian, 1993).

25

30

Telomeres fulfil another important task in association with the replication of eukaryotic cell DNA. In contrast to the circular genomes of prokaryotes, the linear chromosomes of eukaryotes cannot be completely replicated by the DNA polymerase complex. RNA primers are required to initiate DNA replication. After elimination of the RNA primers, extension of the Okazaki fragments and subsequent ligation, the newly synthesized DNA strand lacks the 5' end since the RNA primer cannot be replaced by DNA at that point. Without special protective mechanisms, the chromosomes would therefore shrink with each cell division ("end-replication problem"; Harley *et al.*, 1990). The non-coding telomere sequences presumably constitute a buffer zone for preventing the loss of genes (Sandell and Zakian, 1993).

001260.3428560

In addition to this, telomeres also play an import role in regulating cell ageing (Olovnikov, 1973). Human somatic cells exhibit a limited capacity for replication in culture; after a certain period of time, they become senescent. In this state, the cells no longer divide even after having been stimulated with growth factors; however, they do not die and remain metabolically active (Goldstein, 1990). Various observations support the hypothesis that a cell determines how many more times it can divide on the basis of the length of its telomeres (Allsopp *et al.*, 1992).

In summary, the telomeres consequently possess key functions in the ageing of cells, and in stabilizing the genetic material and preventing cancer.

The enzyme telomerase synthesizes the telomeres

As described above, organisms which possess linear chromosomes can only replicate their genome incompletely in the absence of a special protective mechanism. Most eukaryotes use a special enzyme, i.e. telomerase, for regenerating the telomere sequences. Telomerase is expressed constitutively in the single-cell organisms which have so far been investigated. On the other hand, telomerase activity has only been measured in humans in germ cells and tumour cells, whereas neighbouring somatic tissue did not contain any telomerase (Kim *et al.*, 1994).

Telomerase can also be designated functionally as terminal telomere transferase, which is located in the cell nucleus as a multiprotein complex. While the RNA moiety of human telomerase has been known for a relatively long period of time (Feng *et al.*, 1995), the catalytic subunit of this enzyme group was recently identified in a variety of organisms (Lingner *et al.*, 1997; cf. our application PCT EP/98/03468 which is likewise pending). These catalytic subunits of telomerase are strikingly homologous both among themselves and in relation to all previously known reverse transcriptases.

WO 98/14592 also describes nucleic acid and amino acid sequences of the catalytic telomerase subunit.

Activation of telomerase in human tumours

It was originally only possible to demonstrate telomerase activity in humans in germ line cells and not in normal somatic cells (Hastie *et al.*, 1990; Kim *et al.*, 1994). Following the development of a more sensitive detection method (Kim *et al.*, 1994), a low telomerase activity was also detected in hematopoietic cells (Broccoli *et al.*, 1995; Counter *et al.*, 1995; Hiyama *et al.*, 1995). It is true, however, that these cells nevertheless exhibited a reduction in the telomeres (Vaziri *et al.*, 1994; Counter *et al.*, 1995). It has still not been resolved whether the quantity of enzyme in these cells is not sufficient for compensating the telomere loss or whether the telomerase activity which is measured stems from a subpopulation, e.g. incompletely differentiated CD34⁺38⁺ precursor cells (Hiyama *et al.*, 1995). In order to resolve this, it would be necessary to detect telomerase activity in a single cell.

Interestingly, however, significant telomerase activity was detected in a large number of the tumour tissues which had thus far been tested (1734/2031, 85%; Shay, 1997), whereas no activity was found in normal somatic tissue (1/196, <1%, Shay, 1997). In addition various investigations have shown that the telomeres still shrank in senescent cells which were transformed with viral oncoproteins and it was only possible to detect telomerase in the subpopulation which survived the growth crisis (Counter *et al.*, 1992). The telomeres were also stable in these immortalized cells. (Counter *et al.*, 1992). Similar findings from investigations in mice (Blasco *et al.*, 1996) support the assumption that reactivation of the telomerase is a late event in tumorigenesis.

Based on these results, a "telomerase hypothesis" was developed which links the loss of telomere sequences and cell ageing with telomerase activity and the development of cancer. In long-lived species such as humans, the shrinking of the telomeres can be regarded as being a mechanism for suppressing tumours. Differentiated cells which do not contain any telomerase cease their cell division at a particular telomere length. If such a cell mutates, it can only form a tumour if the cell can extend its telomeres.

Otherwise, the cell would continue to lose telomere sequences until its chromosomes became unstable and it was finally destroyed. Telomerase reactivation is presumably the main mechanism used by tumour cells to stabilize their telomeres.

5 It follows from these observations and considerations that it should be possible to treat tumours by inhibiting the telomerase. Conventional cancer therapies using cytostatic agents or short-wave radiation damage all the dividing cells in the body in addition to the tumour cells. However, since only germ line cells, apart from tumour cells, contain significant telomerase activity, telomerase inhibitors would attack the
10 tumour cells more specifically and consequently elicit fewer undesirable side effects. Telomerase activity has been detected in all the tumour tissues which have so far been tested, which means that these therapeutic agents could be employed against all types of cancer. The effect of telomerase inhibitors would then set in when the telomeres of the cells had shortened to such an extent that the genome became
15 unstable. Since tumour cells usually possess telomeres which are shorter than those of normal somatic cells, cancer cells would be the first to be eliminated by the telomerase inhibitors. By contrast, cells possessing long telomeres, such as the germ cells, would only be damaged at a much later date. Telomerase inhibitors consequently represent a potential way forward in the treatment of cancer.

20 It becomes possible to obtain unambiguous answers to the question of the nature and points of attack of physiological telomerase inhibitors once the manner in which expression of the telomerase gene is regulated has also been identified.

25 Regulation of gene expression in eukaryotes

There are a large number of points in eukaryotic gene expression, i.e. the cellular flow of information from the DNA to the protein by way of the RNA, at which regulatory mechanisms can exert an effect. Examples of individual control steps are
30 gene amplification, the recombination of gene loci, chromatin structure, DNA methylation, transcription, post-transcriptional modifications of mRNA, mRNA transport, translation and post-translational modifications of proteins. Studies which

have been carried out to date indicate that control at the level of transcription initiation is of the greatest importance (Latchman, 1991).

A region which is responsible for regulating transcription, and which is designated the promoter region, is located directly upstream of the transcription start of a gene which is transcribed by RNA polymerase II. Comparison of the nucleotide sequences of promoter regions from a large number of known genes shows that particular sequence motifs occur regularly in this region. These elements include, inter alia, the TATA box, the CCAAT box and the GC box, which elements are recognized by specific proteins. The TATA box, which is located about 30 nucleotides upstream of the transcription start, is, for example, recognized by the TFIID subunit TBP ("TATA box-binding protein"), whereas particular GC-rich sequence segments are specifically bound by the transcription factor Spl ("specificity protein I").

The promoter can be functionally subdivided into a regulatory segment and a constitutive segment (Latchman, 1991). The constitutive control region comprises the so-called core promoter which enables transcription to be initiated correctly. This promoter contains the sequence elements which are described as UPE's (upstream promoter elements) which are necessary for efficient transcription. The regulatory control segments, which can be interlaced with the UPE's, possess sequence elements which can be involved in the signal-dependent regulation of transcription by hormones, growth factors, etc. They impart tissue-specific or cell-specific promoter properties.

DNA segments which are able to exert an influence on gene expression over relatively large distances are a characteristic feature of eukaryotic genes. These elements can be located upstream or downstream of a transcription unit, or within the unit, and can perform their function independently of their orientation. These sequence segments may reinforce (enhancers) or attenuate (silencers) promoter activity. In a similar way to the promoter regions, enhancers and silencers also accommodate several binding sites for transcription factors.

The invention relates to the DNA sequences from the 5'-flanking region of the gene for the catalytically active human telomerase subunit and intron sequences for this gene.

- 5 The invention particularly relates to the 5'-flanking regulatory DNA sequence which contains the promoter DNA sequence for the gene for the human catalytic telomerase subunit, as depicted in Fig. 10 (SEQ ID NO 3).

- 10 The invention furthermore relates to part regions of the 5'-flanking regulatory DNA sequence, as depicted in Fig. 4 (SEQ ID NO 1), which has a regulatory effect.

- 15 Intron sequences for the gene for the human catalytic telomerase subunit, in particular those sequences which have a regulatory effect, are also part of the subject-matter of the present invention. The intron sequences according to the invention are described in detail in the context of Example 5 (cf. SEQ ID NO 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 and 20).

- 20 The invention furthermore relates to a recombinant construct which comprises the DNA sequences according to the invention, in particular the 5'-flanking DNA sequence of the gene for the human catalytic telomerase subunit, or part regions thereof.

- 25 Preference is given to recombinant constructs which, in addition to the DNA sequences according to the invention, in particular the 5'-flanking DNA sequence of the gene for the human catalytic telomerase subunit, or part regions thereof, also contain one or more additional DNA sequences which encode polypeptides or proteins.

- 30 According to a particularly preferred embodiment, these additional DNA sequences encode antineoplastic proteins.

Particular preference is given to those antineoplastic proteins which inhibit angiogenesis directly or indirectly. Examples of these proteins are:

- 5 Plasminogen activator inhibitor (PAI-1), PAI-2, PAI-3, angiostatin, endostatin, platelet factor 4, TIMP-1, TIMP-2, TIMP-3 and leukaemia inhibitory factor (LIF).

Antineoplastic proteins which have a direct or indirect cytostatic effect on tumours are likewise particularly preferred. These proteins include, in particular:

- 10 perforin, granzyme, IL-2, IL-4, IL-12, interferons, such as IFN- α , IFN- β and IFN- γ , TNF, TNF- α , TNF- β , oncostatin M; tumour suppressor genes, such as p53, retinoblastoma.

- 15 Particular preference is furthermore given to antineoplastic proteins which, where appropriate in addition to their antineoplastic effect, stimulate inflammations and thereby contribute to the elimination of tumour cells. Examples of these proteins are:

- 20 RANTES, monocyte chemotactic and activating factor (MCAF), IL-8, macrophage inflammatory protein (MIP-1 α , - β), neutrophil activating protein-2 (NAP-2), IL-3, IL-5, human leukaemia inhibitory factor (LIF), IL-7, IL-11, IL-13, GM-CSF, G-CSF and M-CSF.

- 25 Particular preference is furthermore given to antineoplastic proteins which, due to their action as enzymes, are able to convert precursors of an antineoplastic active compound into an antineoplastic active compound. Examples of these enzymes are:

- 30 herpes simplex virus thymidine kinase, varicella zoster virus thymidine kinase, bacterial nitroreductase, bacterial β -glucuronidase, plant β -glucuronidase from *Secale cereale*, human glucuronidase, human carboxypeptidase, bacterial carboxypeptidase, bacterial β -lactamase, bacterial cytosine deaminase, human catalase and/or phosphatase, human alkaline phosphatase, type 5 acid phosphatase, human

The abovementioned recombinant constructs can also contain DNA sequences which encode factor VIII or factor IX, or part fragments thereof. These DNA sequences also include other blood clotting factors.

The abovementioned recombinant constructs can also contain DNA sequences which encode a reporter protein. Examples of these reporter proteins are:

Chloramphenicol acetyl transferase (CAT), glow-worm luciferase (LUC), β -galactosidase (β -Gal), secreted alkaline phosphatase (SEAP), human growth hormone (hGH), β -glucuronidase (GUS), green-fluorescing protein (GFP), and all the variants derived therefrom, aquarin and obelin.

Recombinant constructs according to the invention can also contain DNA which encodes the human catalytic telomerase subunit and its variants and fragments in the antisense orientation. Where appropriate, these constructs can also contain other protein subunits of the human telomerase and the telomerase RNA component in the antisense orientation.

The recombinant constructs can, in addition to the DNA which encodes the human catalytic telomerase subunit, and its variants and fragments, also contain other protein subunits of the human telomerase and the telomerase RNA component.

The invention furthermore relates to a vector which contains the abovementioned DNA sequences according to the invention, in particular the 5'-flanking DNA sequences and also one or more of the other DNA sequences mentioned above.

The preferred vector for these constructs is a virus, for example a retrovirus, an adenovirus, an adeno-associated virus, a herpes simplex virus, a vaccinia virus, a lentiviral virus, a Sindbis virus and a Semliki forest virus.

Preference is also given to using plasmids as vectors.

5 The invention furthermore relates to pharmaceutical preparations which comprise recombinant constructs or vectors according to the invention; for example a preparation in a colloidal dispersion system.

Examples of suitable colloidal dispersion systems are liposomes or polylysine ligands.

10

The preparations of the constructs or vectors according to the invention in colloidal dispersion systems can be supplemented with a ligand which binds to the membrane structures of tumour cells. Such a ligand can, for example, be attached to the construct or the vector or else be a component of the liposome structure.

15

Suitable ligands are, in particular, polyclonal or monoclonal antibodies, or antibody fragments thereof, which bind, by their variable domains, to the membrane structures of tumour cells, or substances carrying mannose terminally, cytokines or growth factors, or fragments or part sequences thereof, which bind to receptors on tumour cells.

20

Examples of corresponding membrane structures are receptors for a cytokine or a growth factor, such as IL-1, EGF, PDGF, VEGF, TGF β , insulin or insulin-like growth factor (ILGF), or adhesion molecules, such as SLeX, LFA-1, MAC-1, LECAM-1 or VLA-4, or the mannose-6-phosphate receptor.

25

The present invention includes pharmaceutical preparations which, in addition to the vector constructs according to the invention, can also comprise non-toxic, inert, pharmaceutically suitable excipients. It is possible to conceive of administering (e.g. intravenously, intraarterially, intramuscularly, subcutaneously, intradermally, anally, vaginally, nasally, transdermally, intraperitoneally, as an aerosol or orally) these preparations at the site of a tumour or administering them systemically.

30

The vector constructs according to the invention can be employed in gene therapy.

5 The invention furthermore relates to a recombinant host cell, in particular a recombinant eukaryotic host cell, which harbours the above-described constructs or vectors.

10 The invention furthermore relates to a process for identifying substances which affect the promoter activity, silencer activity or enhancer activity of the catalytic telomerase subunit, with this process comprising the following steps:

15 A. adding a candidate substance to a host cell which harbours the regulatory DNA sequence according to the invention, in particular the 5'-flanking regulatory DNA sequence for the gene for the human catalytic telomerase subunit, or a part region thereof which has a regulatory effect, which sequence or part region is functionally linked to a reporter gene, and

B. measuring the effect of the substance on expression of the reporter gene.

20 The process can be employed for identifying substances which increase the promoter activity, silencer activity or enhancer activity of the catalytic telomerase subunit.

25 The process can furthermore be employed for identifying substances which inhibit the promoter activity, silencer activity or enhancer activator of the catalytic telomerase subunit.

30 The invention furthermore relates to a process for identifying factors which bind specifically to fragments of the DNA fragments according to the invention, in particular the 5'-flanking regulatory DNA sequence of the catalytic telomerase subunit. This method comprises screening an expression cDNA library using the above-described DNA sequence, or subfragments of widely differing length, as the probe.

The above-described constructs or vectors can also be used for preparing transgenic animals.

5 The invention furthermore relates to a process for detecting telomerase-associated conditions in a patient, which process comprises the following steps:

10 A. incubating a construct or vector, which contains the DNA sequence according to the invention, in particular the 5'-flanking regulatory DNA sequence for the gene for the human catalytic telomerase subunit, or a part region thereof having a regulatory effect, and a reporter gene, with body fluids or cell samples,

15 B. detecting the activity of the reporter gene in order to obtain a diagnostic value; and

20 C. comparing the diagnostic value with standard values for the reporter gene construct in standardized normal cells or body fluids of the same type as the test sample;

The detection of diagnostic values which are higher or lower than the standard comparative values indicates a telomerase-associated condition, which in turn indicates a pathogenic condition.

25 Explanation of the figures:

Fig. 1: Southern blot analysis using genomic DNA from various species

30 A: Photograph of an ethidium bromide-stained 0.7% agarose gel containing approximately 4 µg of Eco RI-cut genomic DNA. Track 1 contains Hind III-cut λ DNA as size markers (23.5, 9.4, 6.7, 4.4, 2.3, 2.0 and 0.6 kb). Tracks 2 to 10 contain human, rhesus monkey, Sprague

Dawley rat, BALB/c mouse, dog, bovine, rabbit, chicken and yeast (*Saccharomyces cerevisiae*) genomic DNA.

5 B: Autoradiogram, corresponding to Fig.1 A, of a Southern blot analysis in which radioactively labelled hTC-cDNA probe of about 720 bp in length is used for the hybridization.

Fig. 2: Restriction analysis of the recombinant λ DNA of the phage clone P12, which hybridizes with a probe from the 5' region of the hTC cDNA.

10 The figure shows a photograph of an ethidium bromide-stained 0.4% agarose gel. Tracks 1 and 2 contain Eco RI/Hind III-cut λ DNA and a 1 kb ladder from Gibco as size markers. Tracks 3 - 7 each contain 250 ng of the DNA from the recombinant phage which has been cut with Bam HI (track 3), Eco RI (track 4), Sal I (track 5), Xho I (track 6) and Sac I (track 7). The arrows mark the two λ arms of the vector EMBL3 Sp6/T7.

15 Fig. 3: Restriction analysis and Southern blot analysis of the recombinant λ DNA of the phage clone which hybridizes with a probe from the 5' region of the hTC cDNA.

20 A: The figure shows a photograph of an ethidium bromide-stained 0.8% agarose gel. Tracks 1 and 15 contain a 1 kb ladder from Gibco as size markers. Tracks 2 to 14 each contain 250 ng of cut λ DNA from the recombinant phage clone. The following enzymes were employed: track 2: Sac I, track 3: Xho I, track 4: Xho I, Xba I, track 5: Sac I, Xho I, track 6: Sal I, Xho I, Xba I, track 7: Sac I, Xho I, Xba I, track 8: Sac I, Sal I, Xba I, track 9: Sac I, Sal I, BamHI, track 10: Sac I, Sal I, Xho I, track 11: Not I, track 12: Sma I, track 13: empty, track 14: not digested.

30

B: Autoradiogram, corresponding to Fig. 3 A, of a Southern blot analysis. A 5'-hTC cDNA fragment of about 420 bp in length was used as the probe for the hybridization.

5 Fig. 4: Partial DNA sequence of the 5'-flanking region and of the promoter of the gene for the human catalytic telomerase subunit. The ATG start codon in the sequence is printed in bold. The depicted sequence corresponds to SEQ ID NO 1.

10 Fig. 5: Use of primer extension analysis to identify the transcription start.

The figure shows an autoradiogram of a denaturing polyacrylamide gel which was selected for depicting a primer extension analysis. An oligonucleotide having the sequence

15 5'GTTAAGTTGTAGCTTACACTGGTTCTC 3' was used as the primer. The primer extension reaction was loaded in track 1. Tracks G, A, T and C constitute the sequence reactions using the same primer and the corresponding dideoxynucleotides. The thick arrow marks the main transcription start while the thin arrows point to three subsidiary transcription start points.

20

Fig. 6: cDNA sequence of the human catalytic telomerase subunit (hTC; cf. our pending application PCT/EP/98/03468). The depicted sequence corresponds to SEQ ID NO 2.

25 Fig. 7: Structural organization and restriction map of the human hTC gene and its 5'-flanking and 3'-flanking regions.

30 Exons are shown as consecutively numbered rectangles which are filled-in in black, and introns are shown as regions which are not filled in. Untranslated sequence segments in the exons are hatched. Translation starts in exon 1 and ends in exon 16. Restriction enzyme cleavage sites

are marked as follows: S, SacI; X, XhoI. The relative arrangement of the five phage clones (P2, P3, P5, P12, P17), and of the product from the genome walking, are shown by thin lines. As the dots indicate, the sequence of intron 16 has only been partly deciphered.

Fig. 8: HTL splice variants.

A: Diagrammatic structure of the hTC mRNA splice variants. The complete hTC mRNA is depicted as a rectangle with a grey background in the upper region of the figure. The 16 exons are depicted in accordance with their size. The translation start (ATG) and the stop codon, and also the telomerase-specific T motif, and the seven RT motifs, are all shown. The hTC variants are subdivided into deletion and insertion variants. The missing exon sequences are marked in the deletions. The insertions are shown by additional white rectangles. The sizes and origins of the inserted sequences are given. Newly formed stop codons are marked. The size of the insertion in variant INS2 is unknown.

B: Exon-intron transitions in the hTC splice variants. Unspliced 5'-flanking and 3'-flanking sequences are shown as white rectangles. The origins of the exon and intron sequences are given. Intron and exon sequences are shown in small letters and large letters, respectively. The donor and acceptor sequences in the splice sites are underlaid as grey rectangles, and their exon and intron origins are also given.

Fig. 9: Identification of the transcription start by means of RT-PCR analysis.

The RT-PCR was carried out using a cDNA library prepared from HL 60 cells and genomic DNA as the positive control. A common 3' primer hybridizes to a region of the exon 1 sequence. The positions of the different 5' primers in the coding region or the 5'-flanking region are given. In the negative control, no template DNA was added to the PCR reaction. M: DNA size marker.

Fig. 10: Nucleotide sequence and structural features of the hTC promoter.

The figure depicts 11273 bp of the 5'-flanking hTC gene sequence, beginning with the translation start codon ATG (+1). The putative region of the translation start is underlined. Possible regulatory sequence segments within the 4000 bp upstream of the translation start are ringed. The depicted sequence corresponds to SEQ ID NO 3.

Fig. 11: Activity of the hTC promoter in HEK-293 cells.

The first 5000 bp of the 5'-flanking hTC gene region are shown diagrammatically in the upper part of the figure. The ATG start codon is picked out. CpG-rich islands are marked by grey rectangles. The sizes of the hTC promoter-luciferase construct are shown on the left-hand side of the figure. The promoterless pGL2 basic construct and the SV40 promoter construct pGL2-Pro were used as controls in each transfection. The relative luciferase activities of the different promoter constructs in HEK cells are shown as continuous bars on the right-hand side of the figure. The standard deviation is indicated. The numerical values represent the average of two independent experiments which were carried out in duplicate.

Tab. 1: Exon-intron transitions in the hTC gene

The table lists the nucleotide sequences at the 3' and 5' splice transitions of the hTC gene. The consensus sequences for donor and acceptor sequences (AG and GT) are underlaid with grey rectangles. The table shows the intron sequences (small letters) and exon sequences (large letters) which flank the splice acceptor and donor sites. The sizes of the exons and introns are given in bp.

Tab. 2: Potential binding sites for DNA-binding factors in the nucleotide sequence of intron 2

09582246:092100

The search for possible DNA-binding factors (e.g. transcription factors) was carried out using the "find pattern" algorithm from the Genetics Computer Group (Madison, USA) GCG sequence analysis program package. The table lists the abbreviations of the DNA-binding factors which were identified and their location in intron 2.

Tab. I

3' Acceptor Sequence				5' Donor Sequence			
Intron	Exon	Exon	bp	Intron	Exon	Intron	bp
			No.				No.
5' flanking region							
cagggggtctccccgtag catgtccctctctgtttaag gagggtctctctatgtgtag cccatgtgtccccgtag ctcgcctccactccacagag acctctctctctgcggtag ctcccgctctgtcttctgtag ctgtgtctctccgccttag gtattttctctatttttag cattgcctctctgcttag atccccctgtgtcttag tctttctggcgactcttag ctgtccgcctctctcttag agctctctgttttccccctag tctgtattttggccccgtag	GTTTCAGGCACGGTGGGT	GTGCTCTGGCTGAAGAGC	1	CGCCCCCTCTCTTCGGCCAG	gtaggcctccccgggtgag	1	104
	GGTGTGGCTGTGTTCGGGC	GGTGTGGCTGTGTTCGGGC	2	TGGCTGGCAGAGCCAC	gtaggaggtgtgtgacgt	2	8616
	ACACGACGTTGAAGAGGTG	GGTGTGGCTGTGTTCGGGC	3	TGCAAGACATTCGAATCAG	gtaetgtatccccacgcca	3	2089
	GCCGAGCGGTCACTCCGCA	GGTGTGGCTGTGTTCGGGC	4	GTTCCGCGACAGAAAGAGG	gtgctgtgctttgtttta	4	587
	GTGGATGTGAAGCGCGGT	GGTGTGGCTGTGTTCGGGC	5	TGAGCTGTACTTTTGTCAAG	gtgggtgcgggggacccc	5	494
	GTCTCTACCTTGACAGAC	GGTGTGGCTGTGTTCGGGC	6	CAAGCGCTTCAAGAGCCAC	gtaaggttcaactgtgata	6	>4660
	AGCTCTTCCTGGAATGAG	GGTGTGGCTGTGTTCGGGC	7	TGCGGTGGTCAATCGAGCAG	gtttgggcaactgacctgca	7	980
	GTCTAGCTCCAGTGCAC	GGTGTGGCTGTGTTCGGGC	8	CGTGGCCATCAGGGCGCA	gttagtcaggtggccaggt	8	2485
	GGTGTCTCTGGGTTCGGT	GGTGTGGCTGTGTTCGGGC	9	CGGGGATTCGGGGGACG	gttaggctctctctctccc	9	1984
	GACCTGTGGTCGAGGTGTC	GGTGTGGCTGTGTTCGGGC	10	ACGGGAAACCTTCTCTCAG	gttaggcccgtgcccgtgtg	10	1871
	CTATGCCCGGACCTCCATC	GGTGTGGCTGTGTTCGGGC	11	TGCAAGACGACATCTCCAG	gttaggcacactggccgga	11	3801
	GTGAACAGCCTCCAGACG	GGTGTGGCTGTGTTCGGGC	12	CCTGTTTCTGCAATTCAG	gttagcaggtgatgtgta	12	880
	GTGACGAGCTCCAGACG	GGTGTGGCTGTGTTCGGGC	13	TGCTGCTGCAAGCGTACAG	gttagcscgaccacagggg	13	3187
	GTTCACGATGTGTGCTG	GGTGTGGCTGTGTTCGGGC	14	CTGAAGCCACAGAACGAC	gtatgtgacgtgtctggc	14	781
	GGATGTGGTGGGGGCCAA	GGTGTGGCTGTGTTCGGGC	15	CTGGGGTCACTCAGGACAG	gcaagtggtgtgagggc	15	536
	CCGACGAGCAGCTGACTG	GGTGTGGCTGTGTTCGGGC	16	TTTTTCAGTTTGTGAAAAA	3' flanking region		

Tab. 2

Factors	Location in intron 2
C/EBP	2925
CRE.2	2749
Sp1	2378, 4094, 4526, 4787, 4835, 4995
AP-2 CS3	5099
AP-2 CS4	2213, 3699, 4667, 5878, 5938, 6059, 6180, 6496
AP-2 CS5	5350, 5798, 5880, 5940, 6061, 6182, 6375, 6498
PEA3	934, 2505
P53	2125
GR uteroglobin	848, 1487, 2956
PR uteroglobin	3331
Zeste-white	1577, 1619, 1703, 1745, 1787, 1829, 1871, 1913, 1955, 1997, 2039, 2081, 3518, 3709, 4765, 5014, 5055
GRE	846
MyoD-MCK right site/rev	447, 509, 558, 1370, 1595, 1900, 2028, 2099, 4557
MyoD-MCK left site	108, 118, 453, 1566, 1608, 1692, 1734, 1818, 1902, 1986, 2372, 2460, 2720, 3491, 5030
Ets-1 CS	6408
AP1	3784, 4406
CREB	2801
GATA-1	839, 1390, 3154
c-Myc	108, 118, 453, 1566, 1608, 1692, 1734, 1818, 1902, 1986, 2372, 2460, 2720, 3491, 5030
CACCC site	991
CCAAT site	1224
CCAC box	992
CAAT site	463, 2395
Rb site	992, 4663
TATA	3650
CDEI	106, 1564, 1606, 1690, 1732, 1816, 1900, 1984

001260 54228350

Examples

The human gene for the catalytic telomerase subunit (ghTC), and the regions of this gene located 5' and 3', were cloned, while the start point for transcription was determined, potential binding sites for DNA-binding proteins were identified and active promoter fragments were highlighted. The sequence of the hTC cDNA (Fig. 6) has already been reported in our application PCT/EP/98/03468, which is also pending. Unless otherwise mentioned, all the data refer to the position of the cDNA in this sequence.

Example 1

A genomic Southern blot analysis was used to determine whether ghTC constitutes a single gene in the human genome or whether there exist several loci for the hTC gene and possibly also ghTC pseudogenes.

In order to do this, a commercially available zoo blot from Clontech was subjected to Southern blot analysis. This blot contains 4 µg of Eco RI-cut genomic DNA from nine different species (human, monkey, rat, mouse, dog, bovine, rabbit, chicken and yeast). With the exception of yeast, chicken and human, the DNA was isolated from kidney tissue. The human genomic DNA was isolated from placenta and the chicken genomic DNA was purified from liver tissue. An hTC cDNA fragment of about 720 bp in length, which was isolated from hTC cDNA, variant Del2 (position 1685 to 2349 plus 2531 to 2590 in Fig. 6 [deletion 2; cf. Example 5 in Fig. 8]), was used as the radioactively labelled probe in the autoradiogram in Fig. 1. The experimental conditions for the blot hybridization and washing steps were taken from Ausubel *et al.* (1987).

In the case of the human DNA, the probe recognizes two specific DNA fragments. The smaller Eco RI fragment, of from about 1.5 to 1.8 kb in length, probably originates from two Eco RI cleavage sites in an intron in the ghTC DNA. On the

basis of this result, it is to be assumed that only one single ghTC gene is present in the human genome.

Example 2

5

In order to isolate the 5' flanking hTC gene sequence, approx. 1.5×10^6 phages from a human genomic placenta gene library (EMBL 3 SP6/T7 from Clontech, order number HL1067j) were hybridized on nitrocellulose filters (0.45 μm ; from Schleicher and Schuell), in accordance with the manufacturer's instructions, with a radioactively labelled 5'-hTC cDNA fragment of about 500 bp in length (position 839 to 1345 in Fig. 6). The nitrocellulose filters were firstly incubated, at 42°C for two hours, in 2 x SSC (0.3 M NaCl; 0.5 M Tris-HCl, pH 8.0) and then in a prehybridization solution (50% formamide; 5 x SSPE, pH 7.4; 5 x Denhard's solution; 0.25% SDS; 100 μg of herring sperm DNA/ml). For the overnight hybridization, the prehybridization solution was supplemented with 1.5×10^6 cpm of denatured, radioactively labelled probe/ml of solution. Nonspecifically bound radioactive DNA was removed under stringent conditions, i.e. by means of three five-minute steps of washing with 2 x SSC; 0.1% SDS at from 55 to 65°C. The filters were evaluated by autoradiography.

20

The phage clones which were identified in this primary investigation were purified (Ausubel *et al.* (1987)). In subsequent analyses, one phage clone, i.e. P12 turned out to be potentially positive. A λ DNA preparation carried out on this phage (Ausubel *et al.* (1987)), and the subsequent restriction digestion with enzymes which release the genomic insert in fragments, showed that this phage clone contains an insert of approx. 15 kb in the vector (Fig. 2).

25

In order to isolate the complete hTC gene sequence, in each case from 1 to 1.5×10^6 phages were screened, in independent experiments, with in each case different radioactively labelled probes, as described above.

30

The phage clones which were identified in these primary investigations, and which were positive for the corresponding probes, were purified. The phage clone P17 was found to contain an hTC cDNA fragment of about 250 bp in length (position 1787 to 2040 in Fig. 6). The phage clone P2 was identified as containing an hTC cDNA fragment of about 740 bp in length (position 1685 to 2349 plus 2531 to 2607 in Fig. 6 [deletion 2; cf. Example 5]). The phage clones P3 and P5 were found to contain a 3' hTC cDNA fragment of 420 bp in length (position 3047 to 3470 in Fig. 6). After the λ DNA had been prepared from these phages, and subsequently subjected to restriction digestion with enzymes which release the genomic insert in fragments, the inserts were subcloned into plasmids (Example 4).

Example 3

In order to investigate whether the 5' end of the hTC cDNA was also present in the insert in the recombinant phage clone P12, the λ DNA from this clone was hybridized, in a Southern blot analysis, with a radiactively labelled hTC cDNA fragment of about 440 bp in length (position 1 to 440 in Fig. 6) from the extreme 5' region (Fig. 3).

Since the isolated λ DNA from the positive clone also hybridizes with the extreme 5' end of the hTC cDNA, this phage probably also contains the 5' sequence region flanking the ATG start codon.

Example 4

In order to subclone the entire 15 kb insert in the positive phage clone P12 in the form of subfragments, and subsequently to sequence these fragments, restriction endonucleases which, on the one hand, release the entire insert from EMBL3 Sp6/T7 (cf. Example 2) and, in addition, cut within the insert, were selected for digesting the DNA.

In all, two Xho I subfragments, of about 8.3 and about 6.5 kb in length, respectively, and three Sac I subfragments, of about 8.5, about 3.5 and about 3 kb in length, respectively, were subcloned into the pBluescript KS(+) vector (from Stratagene). The 5123 bp 5'-flanking nucleotide sequence of the ghTC gene region, starting from the ATG start codon, was determined by analysing the sequences of these fragments (Fig. 4; corresponding to SEQ ID NO 1). Fig. 4 depicts the first 5123 bp (starting from the ATG start codon). Fig. 10 depicts the entire cloned 5' sequence (corresponding to SEQ ID NO 3).

In order to subclone the entire insert, of approx. 14.6 kb in size, in phage clone P17 in the form of subfragments, restriction endonucleases which, on the one hand, release the entire insert from EMLB3 Sp6/T7 and, in addition, cut a few times within the insert, were selected for digesting the DNA. Three XhoI/BamHI fragments, of 7.1 kb, 4.2 kb and 1.5 kb in size, respectively, and one BamHI fragment, of 1.8 kb in size, were subcloned by means of using a combination digestion with the enzymes XhoI and BamHI. Combination restriction digestion with the enzymes XhoI and XbaI resulted in a XhoI/XbaI fragment of 6.5 kb in size, and two XhoI fragments, of 6.5 kb and 1.5 kb in size, respectively, being cloned.

Digestion with the restriction enzyme XhoI was used to subclone the insert, of approx. 17.9 kb in size, in phage clone P2 in the form of subfragments. In all, three XhoI subfragments, of 7.5 kb, 6.4 kb and 1.6 kb in length, respectively, were cloned. Four SacI fragments, of 4.8 kb, 3 kb, 2 kb and 1.8 kb in size, respectively, were additionally subcloned by digesting with the restriction enzyme SacI.

The insert, of approx. 13.5 kb in size, in phage clone P3 was subcloned by digesting with the restriction enzymes SacI and/or XhoI. Six SacI subfragments, of 3.2 kb, 2 kb, 0.9 kb, 0.8 kb, 0.65 kb and 0.5 kb in length, respectively, and two XhoI subfragments, of 6.5 kb and 4.3 kb in length, respectively, were obtained in this connection.

The insert, of approx. 13.2 kb in size, in phage clone P5 was subcloned by digesting with the restriction enzymes *SacI* and/or *XhoI*. In all, *SacI* fragments of 6.5 kb, 3.3 kb, 3.2 kb, 0.8 kb and 0.3 kb in size, and *XhoI* fragments of 7 kb and 3.2 kb in size, were subcloned.

5

In order to clone the hTC genomic sequence region located 3' of phage clone P17 and 5' of phage clone P2, 3 genomic walkings were carried out using the Clontech GenomeWalker™ kits (catalogue number K1803-1) and various combinations of primers. In a final volume of 50 µl, 10 pmol of dNTP mix were added to 1 µl of human GenomeWalker Library HDL (from Clontech), and a PCR reaction was carried out in 1xKlen Taq PCR reaction buffer and 1xAdvantage Klen Taq polymerase mix (from Clontech). 10 pmol of an internal gene-specific primer, and 10 pmol of the adaptor primer AP1 (5'-GTAATACGACTCACTATAGGGC-3'; from Clontech) were added as primers. The PCR was carried out in 3 steps as a touchdown PCR. First of all, denaturation was carried out at 94°C for 20 sec, and the primers were then annealed, and the DNA chain extended, at 72°C for 4 min, over 7 cycles. There then followed 37 cycles in which the DNA was denaturated at 94°C for 20 sec but the subsequent primer extension took place at 67°C for 4 min. In conclusion, there followed a chain extension at 67°C for 4 min. After this first PCR, the PCR product was diluted 1:50. One µl of this dilution was used in a second nested PCR together with 10 pmol of dNTP mix in 1xKlen Taq PCR reaction buffer and 1xAdvantage Klen Taq polymerase mix and also 10 pmol of a nested gene-specific primer and 10 pmol of the nested Marathon Adaptor primers AP2 (5'-ACTATAGGGCACGCGTGGT-3'; from Clontech). The PCR conditions corresponded to the parameters which were selected in the first PCR. As the sole exception, only 5 cycles rather than 7 cycles were selected in the first PCR step and only 24 cycles, instead of 37 cycles, were run in the second PCR step. The products of this nested genomic walking PCR were cloned into the TA Cloning Vector pCRII from InVitrogen.

30

00502245.092100

In the first genomic walking, the gene-specific primer C3K2-GSP1 (5'-GACGTGGCTCTTGAAGGCCTTG-3') and the nested gene-specific primer C3K2-GSP2 (5'-GCCTTCTGGACCACGGCATAACC-3') were used, together with the HDL library 4, and a PCR fragment of 1639 bp in length was obtained. In the second genomic walking, a PCR fragment of 685 bp in length was amplified from the HDL library 4 using the gene-specific primer C3F2 (5'-CGTAGTTGAGCACGCTGAACAGTG-3') and the nested gene-specific primer C3F3 (5'-CCTTCACCTCGAGGTGAGACGCT-3. The third genomic walking mixture, using the gene-specific primer DEL5-GSP1 (5'-GGTGGATGTGACGGGCGCGTACG-3') and the nested gene-specific primer C5K-GSP1 (5'-GGTATGCCGTGGTCCAGAAGGC-3'), led to a 924 bp PCR fragments being cloned from the HDL library 1. In all, 2100 bp of the genomic hTC region located 3' of phage clone P17 were identified using this genomic walking method (see Fig. 7).

The subcloned fragments, and the genomic walking products, were sequenced in single-stranded form. The Lasergene Biocomputing Software (DNASTAR Inc. Madison, Wisconsin, USA) was used to identify overlapping regions and form contigs. In all, 2 large contigs were assembled from the sequences collected from phage clones P12, P17, P2, P3 and P5, and also the sequence data from the genomic walking. Contig 1 consists of sequence data from phage clones P12 and P17 and the sequence data from the genomic walking. Contig 2 was put together from the sequences from phage clones P2, P3 and P5. Overlapping phage clone regions are shown diagrammatically in Fig. 7. The sequence data from the 2 contigs are shown below. The ATG start codon in contig 1 is underlined. The TGA stop codon is underlined in contig 2.

Contig1:

	ACTTGAGCCCT	ANGAGTTCAA	GGCTACGGTG	AGCCATGATT	GCAACACAC	ACGCCAGCCT	TGGTGCAGCA	70
	ATGAGACCCCT	GTCTCAAAA	AAAAAASAAA	AATTGAAATA	ATATAAAGCA	TCCTTCTCGT	GCACAGTGA	140
5	ACAAAACACG	AAATCAACAA	CAAGAGGAAT	TTTGAAAAAT	ATACAAACAC	ATGAAAAATT	ACAACATATAC	210
	TTCTGAATAG	CAAGTCAGTC	AATGAAGAAA	TTTAAAAAGA	AATTGAAAAA	TTTATTTTAG	CAATATAGTC	280
	CGGAAACATA	ACCTCTCAAA	ACCCACGGTA	TACAGCAAAA	CGAGTGCTAA	CGAGGAAGTT	TATAGTCTAT	350
	AGCGCATCTA	TCAAAAAGTA	AGAAAAGCCA	GGCGAGTGAG	CTCATGTGCT	TAATCCAGCT	ATCTTGGGAG	420
	CCAGAGGGGG	GCAGATCGCC	TGAGTCAGGG	AGCTTGCAGC	CAGCTTGACC	ACACAGAGCA	AGCTTCTGCG	490
10	CTATCAAAAA	TACAAATTA	CGTGGCGATG	GTGGACATGT	CCTGTATCTC	CGAGCTACTG	GGAGGTGAG	560
	CGAGGATAAC	CGCTTGAACC	CAGGAGGTGG	AGGTTGCGGT	TGCGCCGATT	TGCGCCATTG	GACTTCCAGT	630
	TGGGTAAACAA	GAGTGAARCC	CTGTCTCAAG	AAAAAASAAA	AAGTGAASAA	ACTTAAANAT	ACAACCTAAT	700
	GTGACCCCTT	AAGAAGATCG	AAAGCAAGAA	GCAAACATAA	CTTAAATTTG	GTAAAAAGAA	AGAAATGATA	770
15	AGAGTCAGAG	CGAAATTAAT	TGAACATGAA	AGATAACATAT	ACMAAGAGAT	ACAAARATTA	AAAGTGTGTT	840
	TTTTGAAGG	ATAACACAA	TTGACAAACC	TTTGCCAGAG	CTAGAGAAAA	AGGAAAAAGG	ACCTAATAAT	910
	ATAAGATCTG	AGATGAAGAA	AGAGACATTA	CAACTGATAC	CACAGAAATT	CAAAAGATTA	CTAGAGTCTA	980
	CTATGAGCAA	CTGTACATCTA	ATAAATTGAA	AAACCTAGAA	AAAAATAGATA	AATTCCTAGA	TGCTATCAAC	1050
	CTACCAAGAT	TGCAACATGA	AGAAATCCAA	AGCCCAACAA	GACCAATTAAC	ATAATCTGGA	TAAAGCCACT	1120
20	ATAAATAAGT	CTCTTAGCAA	AGAGAAAGCC	AGGACCCAAAT	GGCTTCCCTG	CTGGATTTTA	CAATCATATT	1190
	AAAGAGAAT	GAATTTCCAT	CTCTACTCAA	CTATTCTGAA	AAATAGAGGA	AGAATACTCT	CCAACTCAT	1260
	TCTACATGGC	CAGTATTACC	CGTGTCCAA	AACACGACAA	AAACACATCA	AAACACAAAC	AAACAAARA	1330
	CAGAAAGAAA	GAATACATCA	GGCCAAATCA	CCTGATGAAT	ACTGATACAA	AAATCTCCAA	CAAAACACTA	1400
	GAAACAAAGA	TAAACAACAA	CTTCCGAAAG	ATCATCTCAT	GTGATCAAGT	GGGATTTTAT	CCAGGATGGT	1470
25	AGAAGTGGTT	CAACATATGC	AAATCAATCA	ATGTGATACA	TATCTCCCAAC	AAATGAAGAT	ACAAAAACTA	1540
	TATGATTTAT	CGCTTTTATG	CAGAAAAGAG	ATTGTGATAA	TCCTTGACCC	CTTATGATA	AAAGCCCTCA	1610
	AAACACAGG	TATATAGAAA	ACATAACAGC	AGGACACAGT	GGCTCCACCC	TGCTGATCCA	GCACCTCTGG	1680
	AGGCAAGGT	AGGATGATG	CTTGGGCCCA	GGAGTGTGAG	ACTTGCCTGG	GCACACATAT	TAGTATGATA	1750
	CTACAAAAAA	TTTTTTTAAA	AAATTAGGCCA	GGCATGATGG	GATCTGCTGT	TAGTCCACGT	TAGTCTGGAG	1820
30	CTGAGGTGGG	GAGAACTACT	TAAAGCTAGG	AGGTCGAGGG	TGCAGTGAGC	CTAGAGACAT	TCACTGTACT	1890
	CCAGCTTAGA	CACACAGAAA	AGACCCCACT	GANTTAGAAG	AGAGAGAAGG	AGAGAGGAGG	AGAGGAGGAG	1960
	RAAGGAGAGG	GAGGAGAAGG	AGAGAGTGGG	GGAGAGATGG	GAAGGGAGGG	GAATGGAGAA	GGGAGAGTAG	2030
	AGAAGACATG	TAGATCAAGC	ACAGAGGTAG	ACCGAGGTAG	GTATATGAGG	TTATATGAGG	AAAACTGAA	2100
	AGCCTTCTGT	CTAAGATCTG	GAATAGACAA	TCTCACACTG	TCTCACACTG	TATATACATA	TAGTATAGA	2170
	AGCTCTAGAT	AGAGCAATCA	GATTAAGAGAA	AGAAATTAJA	GGCATCCCAA	CTGGAAGAGA	AGAGATCTAA	2240
35	TTATCTCTGT	TGCAAGATAT	ATGATCTTAT	ATCTGGAAAA	GACTTAGAAC	ACCACTAAAA	AACTATATGA	2310
	GGTAAATTT	GGTACAGCAG	CTACAAAAAT	CAATGTGATA	AAATTCAGAT	TATTTCTATA	TTCCACACGC	2380
	GAACAACTTG	AAAAAGAAC	CACAAAAGCA	GCTACARATA	AAATTAACCA	CTGTAGGAAT	AACCAAGAGA	2450
	CTGAAAGTC	TCTACATATG	AAACTATTAAT	ATTTGTGTAA	AGAAATTTGA	AGAGAGACAA	AAAGAGAGAA	2520
40	AGATATTCCA	TGTTCTATAG	TGGGAAGAA	AAATCTGTGT	AAATGTCTCA	TACTACCCAA	AGCAATTTAT	2590
	AAATCTAAGT	GAATCCCTAT	TAAATACTAT	ATGACGTTCT	TCACAGAAAT	AGAAGAAGAA	ATTCTTAAGT	2660
	TGTGACAGAA	CCACAAAGAA	CCCAAGATAG	CCAAAGCTAT	CCTGACCAAA	AAGACACAAA	CTGGAGGACT	2730
	CAGATTACCT	TGACTCAAA	TATACTACAA	AGCTATAGTA	ACCCAAACTA	CACTGTAGCT	GCATATAAAG	2800
	ASATAGAGCA	TGGCCACAGG	GAACAGAAAT	GAGAGTCAAG	AAACAAATCC	ATGCACTTAC	AGTAATCTCA	2870
45	TTTTTGACAA	AGGTGCCAAG	AGCATCTTTT	GGGGAAGAGA	TATCTCTTCT	ATAATATGCT	GCTGGAGGAA	2940
	CTGGATATCC	ATATGCAAAA	TAAACATACT	AGAACTCTGT	CTCTCCACT	ATACAAAAGC	GAATCAAAAT	3010
	GGATGAAGGG	CTTAATATCA	AAACCTCAAA	CTTTGCAACT	ACTAAAAGAA	AACACCCGAG	AAACTCTCCA	3080
	GCAGATTGGA	TGGGGCAAG	ACTTCTTGAG	TAATTTCCCT	CAGGCACAGG	CACCAACAGG	AAAAACAGAC	3150
	AATGGGATCT	ATATCAAGTT	AAAAAGCTTC	TGCCACACAA	AGGAAACAAAT	CAACCAAGAG	AGAGACACAC	3220
50	CCACAGATG	GGAGATATAT	TTTCAAGACT	ATTCATCTAA	CAGAGATATTA	ATACACACTA	TATATAGGA	3290
	GCTCAAACTA	CTCTTATAGA	AAAAACCTTA	ATAAGCTGAT	TTTCAACACT	ATGCAAAAGT	ATGCAAAAGT	3360
	CTCTTCTCAA	AATTAAGTCAT	ACAAATGGCA	AACAGGCATC	TGAAAATGTT	CTCACACCA	CTGATATCTA	3430
	GAGAAATGCA	AATCAATGAT	ACTATGAGAG	ATCATCTCAT	CCAGTTTAAA	ATGGCTTTTA	TTCAAATAGC	3500
55	AGGCAATTAAC	AAATGCCAGT	GAGAGTGGG	ATAAAGAGAA	ACCTTGGAC	ACTGTGTTGT	GGATGGAAA	3570
	TGTGTCACAC	TATATGAGAC	AGTTTGAAGT	TTCTTCAAAA	AACTAAAAAT	AAAGCTACCA	TACAGCAATC	3640
	CGATCTTAG	GTATATACTG	CACAAAAGGG	ATCTGATGTA	ATCAGAAAGT	ATCTCCACTG	CCAAATTTAT	3710
	TGCGACACTG	TTCAATAGCAG	CCAAGGTTTG	GAAAGCAACT	CAGTGTCTAT	CACACAGACA	ATGGAAGAAAT	3780
	AAATATGTGT	GCACATACAC	ATGGAGTATC	TACGCAAGCA	TAAAAAGAA	TGAGATCTG	TCAGTGTGAA	3850
60	CAGCATGGGG	GGCATCTGTC	AGTATGTTAA	TGAAATAGAT	CAGGACAGAG	AAAGACAAAC	TTTTCAATGT	3920
	CTGCCCTTACT	TGTGGGAGCA	AAATTTAAAA	CAATTTGACAT	AGAAATAGAG	GAGAAATGTT	GGTCTTAGGG	3990
	GTGGGGGAC	AGGGTGACTA	GAGTCACAR	TAATTTTGTG	TATGTTTAAA	ATAATCAATA	AGATATTAAT	4060
	TGGGCTGTTT	TTGATGAGCA	TTGATGAGCA	ATGCTGAGTA	TGCTGAGATA	CTGCTGTTAT	CCGCTGTTAT	4130
	TTATTCACCA	TTTATGTGCT	CTACAAAA	CTCTCATGTA	TGCTATAGTA	ATAAACCTTA	CTATATATAA	4200
	AATTAATAAT	TTAATGGCCA	GGCACGGTGG	CTCATGTCCG	TAACTCCAGC	ACTTTGGGAG	GGCGAGGGGG	4270
65	GTGGATCACC	TAGGATCAGG	AGTTTGAAC	CAGCTTGGCG	ACCATGATGC	AGCCCTGTCT	CTACTAAAGA	4340
	TACAAAAATT	AGGCCAGCGT	GGTGGACACA	ACCTGTAGTC	CAACACTACT	AGGAGGAGTA	GACAGAGAGA	4410
	TTCTTTGAAC	CTGGAGAGGG	GAGTGTGCA	TGAGCCGAGG	TACTGCCACT	CGACTGCAGC	CTGGGTGACA	4480
	GGCAAGGCT	CAATCTCAA	ACAAAAGCAA	TAATTAATTT	AGAAATAGTG	ATTAATATTT	TATTTTATAT	4550
	AATATATACT	CTACTATATT	AGAAAGTTAA	AATTAACAA	ATTATATAAG	GTAATTAAC	ACTTAATCTA	4620
70	AATTAAGAAC	TTATATATGT	GGGTTTCTAS	CTTCTGAAGA	AGTAAAGTT	ATGGCCCAAC	TGGCCCAAT	4690
	GTGAGGAGAG	AACAGTGGAA	TGATCTGTTG	TTAGACGCTC	ATACTCTCTG	TAAAGTGTCT	AAATTTAAAC	4760
	AAACACAGGG	TGGGAGAAAG	TAAAGAGCCA	TTCTTATAGC	CCTAABACAA	CTGCTTAATA	TGGTGAAGAG	4830
	TAACTCTGAT	TATATACAGA	ACAAAAGCAA	TAATTAATTT	ATAGAGGCTC	AGATATGCTA	GGTCTCTGAT	4900
	CACCTGCTTC	TACTTCAAGG	TGCTTTTT	CTGTGTGCTG	TGGAGATTTT	CGATTTGTTG	TTCCTTTGTT	4970
	GTTAACCTTA	CTATGTATGA	ATCTGTAAAC	GAAAAATGGT	GGTGATTTCG	TGACAGAGAA	TTAGAGTATC	5040
75	TGGAGGAAG	CAGGTGGCTC	TGTGGGAGCT	AGGCCACTTA	ATCTTCAAGG	GTCTCTGGCG	AMGACCTAGG	5110

00562246-092100

	CCCAAGTCGC	GGGGAAGTGT	TGCAGGAGGG	CACCTCCGGGT	GGTCCCGGCT	GGCCCTCCAG	GGAGCAATGC	10640
	GTCTCTGGGT	TGCTGCCAG	CCGCGTCTAC	CGCGCTCCGT	CTCTCCCTTC	ACGTCGGGCA	TTCGTGGTGC	10710
	CCGAGCGCG	ACGACCGCG	TCCGACCTGT	GAGCAGCGCT	TGGCTCTCCG	GATCGAGCGA	GGCGCAAG	10780
	GGTGCGCCA	CGCACTCTT	CCAGGCGCT	CCACATACCT	GGCCTCCCT	CGAGGTATCC	CACAGCTGAC	10850
5	GCCGATTGGA	CTCTCTCCG	CTGGGGCCCT	CGCTGGCTGT	CTGTCACCTT	GGGAGCGCGA	CGCGCGGCG	10920
	GGCGGGGAA	CCGCGCCAG	ACCCCGGGT	CCGCGCGGAG	CAGCTGCGCT	GGGCGGCGCA	GGCCGCTGCT	10990
	CCAGTGGATT	CCGCGGACA	GACGCCGAG	ACCGCGTCC	CCAGCTGGCG	GAGGAGCTGG	GACCCCGCGG	11060
	ACCGCTGCTT	CCCTCTACG	TTCAGGCTCC	CGCTCCCTCC	GGCGGACCGC	GGCCGCTCCC	GACCCCTCCG	11130
	GGCTCCCGG	CCGCGGCGG	CGCTCTCCCT	CGCTCTCCCT	TCGCGGCGCG	GGCTCTCTCG	GGCTCTCTCG	11200
10	TCGCGCGCG	AGTTTCAGG	ACGCGTGTGT	CTGTCTGCG	ACGTGGGAG	CCCTGCGGCG	GGCCACCCCG	11270
	GGGATGCGG	GGCTCCCGG	CTGCCGAGCG	GTGCGCTCCC	TGCTGGCGAG	CCAGTACCGC	GAGGTGCTGG	11340
	CGCTGGGACG	GTGCTGGGG	CGCTTGGGG	CCAGGGGTGT	CGGCTGGTG	CAGCGCGGGG	ACCCGCGCGG	11410
	TTTCCGCGG	TTGCTGGCC	AGTGGCTGGT	GTGCTGGCT	TGGGAGCGAC	GGCCGCGCCC	GGCCGCGCCC	11480
	TCCTTTCGCG	AGGTGGGCT	CCCGGGGCTG	GGCGTCCGG	TGGGTTTGG	GGGCGGCGGG	GGGAGACGAG	11550
15	GACCTGGGA	GACGACGCA	GGCGACTCAG	GGCGCTTCC	CCGACGTTGT	CTGCTGTGAA	GGAGCTGGTG	11620
	CGCCGAGTC	TGCAGAGCT	GTGCGAGCG	GGCGGAGGA	ACGTGTGAG	CTTCCGGTTC	CGCTGGCTGG	11690
	ACGGGCGCG	CGGGGCGCC	CCCGAGGCT	TCACACCAAG	CGTGGCGAG	TACCTGGCCA	ACACGGGTGAC	11760
	CACGCACTG	CGGGGAGCG	GGCGTGGGG	GCTGTGCTG	CGCGCGTGG	CGGACGACG	CTGTGGTTGAC	11830
20	CTGCTGGAC	CTGTCGCGT	CTTTGTGCTG	GTGGCTCCCA	CTGCGGCTA	CCAGGTGTGG	GGCGCGCGCG	11900
	TGTAGCACT	CGGCGTGC	ACTCAGGCC	GGCCCCCGC	ACAGCTAGT	GAGCCCGGAA	GGCGTCTGGG	11970
	ATGCGAACG	GCTGTGAAC	ATAGCTCAG	GGAGCGCGG	GCTGCCCTGG	GCTGCCACG	CCGGGTGGCG	12040
	ASGAGGAGCG	CGGCGAGTC	CAGCGAAGT	CTGCGTTGC	CCAGAGCGG	CAGCGTGGCG	CTGCTCCCTG	12110
	ACCGGAGCG	GACGCCCTT	GGCGAGGGT	CTTGGGCCA	CCCGGGCAGG	ACGCGTGGAC	CSAGTAGCGG	12180
25	TGGTTCTGT	GTGTTGTAC	CTGCCAGACC	CGCGAGGAA	GGCCTCTTT	TGAGGGTGGG	GCTCTCTCCG	12250
	ACCGCGCACT	CCACCCATC	GTGGGCGCG	CAGCACCGC	CAGCGCCCC	ATCCACATG	CGCGCATCCG	12320
	GTGCTTGGG	ACCGCCTTT	CCCGCGGTT	ACGCGGAGC	CACGCACTT	CTCTACTCT	CAGGCGACAA	12390
	GGAGCGCTG	CGGCGGCTG	TGCTGCTG	TCGCTGCTG	CTGCGCTGG	CTGCGCTGG	GGAGCGGCGG	12460
	GAGCACATCT	TTCTGGTGT	CAGCGCTGG	ATCCGAGGA	CTCCCGCAG	GTTCGCCCGC	GTGCCCGAC	12530
30	GCTACTGGCA	AATCGGCCG	CTGTTTCTGG	AGCTGCTGG	GACACACCG	CAGTGGCCCT	ACGGGTTGCT	12600
	CTCAAGACG	CTGACGCCG	TGGGAGCTG	GGTCAACCG	GAGCGCGTG	TGTGTCGGC	GGAGAGCGCC	12670
	CAGGGCTCTG	CGCGGCGCC	CAGGAGGAG	GACACGAGC	CCCGTGGCT	GTTGCGAGTG	CGGCGCTGAC	12740
	ACAGCAGCTC	CTGGAGGTG	TACGGCTTCG	TGCGGGCTGT	CTGTGGCGG	GTGTGTCGCC	CAGGCTCTG	12810
	GGGCTCCAG	CGGCGGCTG	CTGCTGCTG	CAGGACCTG	AGGAGTTTC	TCTGCTGGG	GGGCGGCGG	12880
35	AGCTCTGCG	TGGGAGCTG	CAGCTGGAG	ATGAGCGTG	GGGACTGCG	TTGCTGCGC	AGAGAGCCAG	12950
	CTAGGAGGT	GTTGGCGCT	GAGGGCCAG	CGCCACGAG	TGAATGAGT	AGGCGGTTCAG	AAAGGGGCGG	13020
	AGCGAGGAG	GTGCTCTCC	TGCTCTCATC	GTCACTGGG	CACAGCTGGC	TTTTCGCTCA	GGAGCTCGAG	13090
	TGAGACCGGT	GATCTGCGC	TCTGCTCTCC	CTCTGTCCA	GTTTGGATGA	TACCTTACGA	TTCACTTACGA	13160
	CGTTTGTATG	GACAGCGGT	TTCAGGAGCG	CGAGCGGAGA	CAGGTGAGCA	GAGGAGGCTG	GAGGAGGCTG	13230
40	GGAGCGCGG	CTGCGGCTG	TGGCGAGAA	TGCGGAGAA	CTGCGGAGG	CTGCGGAGG	TGCTGCTGCG	13300
	TAACTATTA	CTCTCTTGC	ATTTTCAAG	TGGGGAATGA	GAGGTGGGGA	CAGAGAATCC	CTTCTCTGCG	13370
	GGGTGGGAG	TAAAGGTTT	TCAGGTGAC	GTGGTCAGC	AATATGAGG	TTTGTGTGTA	AGATTTATTT	13440
	TGTGTTTAC	GGCGAGGTC	GTTGGCTCAC	CCGCGTAAT	CACGACATTT	GGGAGCTGCA	GAGGAGTGA	13510
	TCACCTGAGG	TGAGGATTT	GAGCAGGCG	TGACACATG	GTTGAAATCC	TATCTGTACT	AAAAATACCA	13580
45	AAATTAGCTG	GACATGGTG	TGTTGTGCTG	TAAATCCAG	TATTTGGAG	GCTGAGGAG	GAGATACAT	13650
	TGACCCGCTG	AGCGGAGCG	TGAGTGGAG	TGAGATTGT	CCATTGTACT	ACACGCTGCG	GGGACAGAGT	13720
	GAACTCTGCT	CTTTAAAAA	AAAAATGTT	CGTTGATTG	GCCAGGACAG	GGTAGAGGGA	GGGAGATAAG	13790
	AGATCTCTCC	AGACACAGAT	CTGTTCCCAT	CTTTAGGTAT	GAGAGGGGCC	ACATGGGACG	AGAGGACAGC	13860
	AGATGGCTCT	ACCTGCTAG	GAGGAGGAC	TGTTTGGGG	TGTTCAAGGG	ATGGTGTCG	TGGGCGCTTG	13930
50	CGTTGTCOCA	CCCTGTTTT	TGAGGATTGA	TGTTGAGGAA	CTCCGCTGT	AGCCCGCTTT	TGGCTCCGAC	14000
	TGCTCCGAG	CTGCTAGCTG	CGAGTAGAGA	CTAGCTGAG	TTTCAACCG	TGCCGACAA	CTCCGACAGC	14070
	ATGTAAAGAT	TCCGCGCATC	CAGACAAAGA	GGGTGACCTT	TTTGGGGCTC	TTTTTTTTCT	TTTTTTTTCT	14140
	TATGTGTGGC	AAAGATCAT	TAAATGAGA	TGGGCACTCC	TACACAGATT	TCTGTGTGAT	GAGTAGAAT	14210
	TGCTAACTCG	CGGGTGTCTA	CAGCAGGTTG	CTTGAATGCG	TGCGCTTTCG	GATCTGGAA	GTCCTTACCC	14280
	ATCGAACGCG	AGCTGCTTTA	CAGCTGTGCG	GGCTCAGGTG	GACACGCGCG	AGTACAGATTA	GGTCACTGGA	14350
55	ACCGAGTTT	GCTTTTTGTT	CTCAGCTG	CTCTGTGAT	GAGATTTGTA	GTCTCTGAT	CAGAGCTGAT	14420
	CTGTCTATG	CTGCTGCTG	ACTCTAGATG	CTGCTGCTG	CTGCTGCTG	CTTTATGATG	GAGTAGGCGG	14490
	TGCTCCCGGG	GTGTCCCTGT	CAGCTGCAGG	GTGATGAGG	GTGTGCCCC	AGGTTGCCCT	GTACAGTGTA	14560
	GGGTGAGTGA	GGCGCGGCC	CCGGGTGTC	CTGTCCGCTG	CAGCGTGATT	GAGGTTGGCG	CCCGGGTGTG	14630
	CGCTGTCAAG	TGAGAGTGA	GTGAGCGGCC	ATCCCGGGT	GTCCCTGTCA	GTCTGTAGGT	GAGTAGAGCG	14700
60	GTGTCGCCGG	GTGTCCCTGT	CCGTCGAGG	GTGATGGAGG	CAGCTGTGCT	GGGTTGTCCCT	GTACAGTGCA	14770
	GGGTGAGTGA	GGCGCGGCC	CTGTGCTGCT	CTGTGAGTGT	TAGGTTGGCT	GAGGCGGGG	CGAGGAGTGT	14840
	CCCTGTCAAG	TGAGGAGTGA	GTGAGGCACC	GTCCCTGAGG	GTCTCTCCA	GATATAGGCT	GAGTAGAGCA	14910
	CTGTCGCCGG	GTGTCCCTGT	CAGCTGCAGG	GTGATGGAGG	CGCGGCCGCC	GGGTTGCCCT	CTCAGGTGCA	14980
65	GGGTGAGTGA	GGCGCTGTCC	TGCGGTGTC	CTGTCTGCTG	TAGGTTGAGT	GAGGCTGTCT	CCCGAGGCTGT	15050
	CTGTTGGCTT	TGCTCACTG	AGCTTGTCTC	TGAATGTTTG	CTTCTTCTAT	AGGCACTAGT	CGCGCGGTTG	15120
	CCGATGGCT	GGGTAGATG	TGAGGCGGCA	CTGCTGTGCT	CACAGCTCAT	CTTCTTGAT	GTGCTGGCTT	15190
	TGTTGTGAC	CTCTGCTGCG	CATTGACTA	CGGCGGCTG	GAGTGTGAG	CTCTGGCTG	CTCTGGCTG	15260
	GGCACTGAC	CCGACAGCTG	AGGTCGCGTT	GGCTGTGTTG	GGCTGGCTT	GCTCACACG	TGCCCGGAC	15330
	ATGACTGCTG	CCAAATCTCC	TCTCCAGCT	TGCTCTATCG	CAGGCTGGA	CTTGGGCTGT	CTCTGGTGTG	15400
70	CTGCCAGCTG	TGCTTGGAGA	CATCCAGAA	AGGGTCTCT	GTGCCCTGAA	GGAAAGCAAG	TACCCCGGAC	15470
	CCGCTCACTT	GTGCTTTTTT	CTCGAAGCT	GGGCTGTGCT	TGGCGGGTGG	TGGCGGGTGG	CGAACGCTTG	15540
	TACCTTATT	CTGGGACACT	GGCGGCTATT	GCTTAGGCTG	GGCTGTGCT	CGCTGTGCT	CTCTACGCTG	15610
	TATTAAGTGC	ACGACAGCTG	GGGCTGTGCT	CTGTGCTG	GAGGCTGTG	GAGGCTGTG	GAGGCTGTG	15680
	CTGTCGCCCA	GCTCTGCTCA	GACTTCTCTC	GTTTGTGAA	TGACTGATT	ACCTCTGACG	TTAGTTTATG	15750
75	TTTGTATGTC	TGCAATTGAT	GGTTTGTGTT	GGTTTATGTT	TGATTCTCT	TCTGATGCT	TGATTTTATG	15820
	CATGCTGCTG	CTGTAAGTGT	TGCGTTTACC	TGAGTCTGTT	GTGTTGATG	GAGTAATCT	CACATCTCAG	15890
	CAGCTTCAAG	TGTTTGTAAA	TATCTCAAA	GTGTATATC	GTCTTTTAA	TATCTTTAT	CTGTGATTGT	15960
	TTTGTGTTGG	CAGCGTGTGT	TGCGTAGGA	TATCATGGA	ATCATGATGA	CTTTTAGTA	TCTTTAGCT	16030

	TATTCCTGTA	TTTCTTTGAG	CAGTGAAGTA	TTTGAACACT	GTTTATGTTT	AAGATATGTA	GAGTATCAAG	16100
	ATACTGTAGAG	TATTTTAAAG	TATGACTTTA	TATTTAGTTT	CTAACTCAGT	TGGTATGAGT	TCTGTATAAAT	16170
	ACCAATATAT	TGAAGTTTGC	GATGACCTTTG	TTTGTGATCT	AGTGTGTGCA	TGGTITTCAG	AGCTGTCCAT	16240
	TGTAAATTTG	ACATGCTGTG	ANTAGTGGGG	ATGCATGTCT	ACTATATGCA	GCTTATTAGG	ACTGTGCAAT	16240
5	AAGCTCTCTG	CTCTCTCTAG	ATCCATGAAA	TTTCCAGAAC	GAGGCCATAG	TCCCTCAGCT	GGGGGATGGG	16380
	TCTGTTCATT	TCCTTCTCGT	TGCTAGCAAT	TATGTGAGGC	ATTGTATAGT	GCATGACAGT	GGTAGAATTT	16450
	TTATCTTCCT	GATGAGCTGAA	CTTTTGGAG	ACTTCTATGT	CTCTAGTAAT	CTAGTAATTC	TTTCTTTAAA	16520
	TGTGCTCTAT	TACTGCCGCA	CTGGGCTTCT	TTTGAATAGT	ATTTTCTCCT	TGTGCTGTGT	TTTTCGCTTT	16590
10	AATTTATATA	TACATATATA	TTTTTTTTTT	TTTTGAGACA	GAGTCTTGCT	CTGTGCGGCA	GGGTGAGTGC	16660
	AGTGTGTGTA	ACTGACGATA	GTGTAACTTT	TACTCTCTGT	CACCTGGGTA	CTCTCTCAGT	CAGGCTCTGT	16730
	AGTGTGTGTA	CCGAGGCTGG	TCTCAAACTC	TTTGACTCAA	GGGATCCATC	TACCTCGGCT	TCCCAAAAGT	16800
	TGCGAATTAAT	CCGATGAGGC	ACCATGTCTG	GCCTAATTTT	CAACACTTTT	ATATTTCTAT	AGTGTGATGA	16940
15	TGTCTCTGTA	ACAGCATATGA	GTTGAATTTG	CAATCCAGTC	TGACAGTGGT	TGTTTAACTG	GATAACCTGA	17010
	TTTATTTTCA	TTTTTTTGTG	ACTAGTAGGC	CGCGTGGTGC	ACTCTGATTC	TCCAGTGGC	TGTTGCAATG	17080
	CTGTGTCCG	TGTTTCTCTA	CAGTCTCTG	GGTGGCATAT	TGCGTTTCTC	CCGAGTGTG	TGTTGATCTC	17150
	CTGTGTCTCT	CTGTGCTAGT	GGGATTTTGC	TTTTTTTCTT	CTTGTCTTAG	TGTTTACCGT	TGATCTTTTT	17220
	ATTGTCTGTT	TTTGTCTTTG	TTTATTTAGA	CAGTCTCACT	CTGTCAACCA	GGCTGGAGTG	TAATGGACCA	17290
20	ATCTCGGCTC	ACTGCAACCT	CTGCTCTCTC	GGTTCAAGCA	GTTCCTATTC	CTCAACCTCA	TGATGATGCT	17360
	GGATTACAGG	CCGCAACAC	CAGCGCTGGC	TAATTTTTGT	ATTTTTAGTA	GAGAATAGGT	TTCACCAATG	17430
	TGGCCAGGCT	GGTCTCAAC	TCTTGACCTC	AAAGTGAATG	CCGCGCTTGG	CTCTCCAGCA	TGCTGGGATT	17500
	ACAGGTGCAA	CCGACCGTGC	CCGCAATACC	TTGATCTTTT	AAATGAAGT	CTGAACCAAT	GATTCAGCTT	17570
25	TCTGAGCAAA	TAAAGCCCTT	AGTGTATTTT	AGCTTGGGCC	ACCGCCGCTG	CTGTGTCTGT	TTCTTCTGTC	17640
	TGACTTATTT	CTATCTCAGG	CATCTTGACA	CCCCCAAG	CTAAGACATA	TTAATATTTG	TTTCTCGTGT	17710
	GAGTGTGAT	TGAGCTTTTG	CCCGCGCTCT	CTTCTTCTCT	TTTGTCTCCC	GTCTGTCTGT	TGTCCTAGGC	17780
	CCCGCGCTCT	GGGTCTCCTT	CTTCTGCTCT	TGCGTGTCTT	TTCTGTCTGT	TATGTGTGCT	TAAACCCAGC	17850
	CTTACTCTGT	CTGCGGCTCC	ATGGCATCTA	CGGACGCTCG	GGGACGCTCT	CTTATGATGC	CAAGACGAAG	17920
30	ATCTGAGAG	TCACGCTGTC	TCAGGCTGAT	TTGCGCTG	GAACGCTG	GCACACGATG	CCGACGTTTC	17990
	CTTAGCAGT	GGAGCAGAG	AAGCTCCGCT	CGGCTGGGCT	CTAGCTGTGA	AAACCCAGG	CATGTCCGAG	18060
	CTGTGTGGCT	CCGCGCTGTC	GAGTTTGAAA	TGCGGCAAC	CTCGGCTGGA	CGGCGAGCTC	TGAGCGTCTC	18130
	GCTCGGGGCG	GAGGTGTCTG	CTTCTCCCTT	CTGCTTGGG	AACCAAGACA	AAGGATGAGG	CTCGGAGGCC	18200
	TGCTGGCCA	ACAGAGAGAT	GAGGTGAGCC	ATGTGGATAA	TTTTAAATTT	TCTAGGCTGT	CCGCGGCTGGC	18270
	TCAAGCTGCT	AATCCCAAGC	CTTTGGAGAG	CGAAGGCGG	TGAGTACAGA	GCTTCAGGAG	TCGAGACCAT	18340
35	CTCGGCAAG	ATGATGAAG	CCCATGTA	CTAAATAG	AAATATAG	TGGGCTGTGT	GGTGTGCTG	18410
	TGTAATCCCA	CTACTCTGGG	AGGCTGAGGC	AGGAGAATTT	CTTGAAGCT	GGAGTGGGA	TTTCCAGTGA	18480
	CGCGACATCT	CACCACTGCA	CTCCAGCTGT	CGAACACAG	GAGACTCTGT	CTCAAAAAAG	AAAAAATAAA	18550
	AAAAAATAAA	AATTTCTAGTA	GCCACATTAA	AAAAAGTAAA	AAGAAAGAAA	GAAATTAATG	TAATTAATGA	18620
	TTTTACTGAA	GGCCAGCATG	TGCCACACTC	ATCATTTTAT	GTGTGTAATG	GTGGGAGCAT	CATCTCAGAG	18690
40	AAATTTGACA	TTTCTTGAGC	CTTGTGAGCT	GTGTGAGCT	TGTAGTGGT	GTCGCTGGCC	ATCTCGGCTC	18760
	GGGCTGCTGT	GGGCTGCTGT	TGCGATGGCT	GTGTGACCA	ATGTGTCAGT	TCGGGGATGA	CTGGGCTGTC	18830
	CCCTCAGTGA	CTGTGATGTC	CAGTGTCCGG	ATGTGTCAC	TGTGGGATGA	GGTCCGACAG	CCCTGCTGTG	18900
	AGCTGATGCT	GTGTGATGTC	GATGTGTCAG	GTCAAGGCTG	AGGCTCTCAG	GCCTCTGGTG	AGCTGAGGAT	18970
45	ATGAGTGTGC	GATGATGACG	GTCGCGGGTG	AGGTCGCGAG	GCCTCTGTCT	GAGCTGAGAT	TGTGTGCTCT	19040
	GGATGTGTGA	GGTCAAGGGT	GAGGTGTCCA	GGGCTCTGGT	AGGCTGAGG	TATGGAGTCT	GGATGATGCA	19110
	GGTCTGGGCT	GAGTGTGCA	GGGCTCTGCT	TGAGTGTGAT	GTGTGTGTC	TGAGTGTGTC	AGGCTGTGGG	19180
	TGAGGTCAAC	AGGCTCTGGC	GTGAGCTGGG	TGTGCGGTGT	CTGAGTGGTG	CAGGCTGGA	GTGAGTCTGC	19250
	CAGACGTGCG	CAGACCATCT	GGTGAGCTGG	ATATGCGGTG	GCAGGTGGTG	GCAGGTGGTG	GGTGAAGGTTG	19320
	CCAGGCGCTG	CTGTGAGTTG	GATGTGGGTT	CTGCGGATGT	TGCAGGCTGT	GTGTGAGGCT	ACCAAGGCCCT	19390
50	CTGTGTAGCT	GTGTGTGTGG	TGTGTGATGT	GTGCAAGTCT	GGGGTGAAGG	TGCGCAAGCC	CTCTGTGTGT	19460
	AGCTGTGATG	GGTGTGTCTG	GATGTGTCAG	GTGCTGCGG	AGGCTGCGAG	GCCTCTGGTG	AGCTGAGTAT	19530
	CGAGTGTGCA	GATGTGTCAG	GAGTGTGGTG	ATGTCGCGAG	ACCTCTGGCT	GAGGCTGAT	GGGCTGCTG	19600
	GGATGTGTGA	GGTGTGAGT	GAGGTGGCCA	GGGCTCGGT	GAGGCTGATG	TATGGAGTCT	GGATGTGGCT	19670
	TGCTCGGGGT	GAGGTGCCCA	GAGGCTGCTG	TGAGTGTGAT	GTGCGGTGTC	TGAGTGTGAT	AGGCTGTGAG	19740
55	TGAGGTGCGC	AGACCTCTGCT	GTGAGCTGGA	TATGCGGTGT	CCGAGTGGTG	CAGGCTCAGG	GAGGCTCTTC	19810
	CAGGCTCTGC	TGAGCTTGA	GGTATGAGT	CGGATGATG	CAGGCTCGGG	GTAGAGGCTG	CAGGCTCTGC	19880
	TGTGAATGAG	ATGTGTCAG	CTGTGATGCT	GGGTTGTGCT	GGGTTGTGCT	CGAGGCTCTG	GTGATGCTG	19950
	AGTATGGAG	TGAGATGATG	CGAGTCTGAG	CGAGCTGCTG	CGAGGCTCTG	CGAGGCTCTG	CGAGGCTCTG	20020
	CTGTGGAATG	TGAGGCTCTG	GGGTTGTGCT	GCGAGGCCCT	GGTGTGCTG	GAGGATGGA	CTGCGGATGA	20090
	TGCAAGTCTG	GGGTGAGTGT	GCGAGGCCCT	GCTGTGAGCT	GGATGTGCTG	TATCGGATG	GTGCAATGCT	20160
60	GGGTGAGTCT	GCGAGGCCCT	GCTGTGAGCT	GGATGTGCTG	TATCGGATG	GTCAGGCTGT	GGGCTGAGCT	20230
	CAGGAGGCCCT	GGGTGTGAGT	TGTTGTGCTG	GTGCTGCTG	GGGTCAGGCT	CGGCTGAGT	CGGAGGCCCT	20300
	CTGCTGAGG	TGCAATCTG	GTGCTGCTG	GTGCTGCTG	GGGTCAGGCT	GGGTCAGGCT	GGGTCAGGCT	20370
	TGCTGGGCTG	GATGCTGCTG	CTGCGATGCT	TGCAAGCTG	GGGTCAGGCT	GGGTCAGGCT	GGGTCAGGCT	20440
	GATGCTGCTG	GTGCTGCTG	TGCAAGCTG	GGGTCAGGCT	GGGTCAGGCT	GGGTCAGGCT	GATGCTGCTG	20510
65	GCTCGGATG	TGCAAGCTG	GGGTCAGGCT	GCGTGAAGCT	GGGTCAGGCT	GGGTCAGGCT	GGGTCAGGCT	20580
	GTGCAAGTCT	GGGTCAGGCT	AGCCAAAGCC	TTGCGTGAAG	TGATGTGAGT	GGGTCAGGCT	GGGTCAGGCT	20650
	CGGCTGAGCT	TGCGCAAGCT	CTGCGTGAAG	GGGTCAGGCT	GGGTCAGGCT	GGGTCAGGCT	CGGCTGAGCT	20720
	GTGCAAGTCT	GGGTCAGGCT	CTGCGTGAAG	GGGTCAGGCT	GGGTCAGGCT	GGGTCAGGCT	GGGTCAGGCT	20790
	CCCTGCTGCT	AGCTGATGAT	CTGCGTGAAG	GATGCTGAG	CTGCGGCTG	AGGTCAGGCT	CCGCTGAGCT	20860
70	GAGCTGAGT	TGCTGTATG	GGATGTGCA	GGTCTGGCT	GAGGTCGCA	GGGTCAGGCT	TGAGTGGAT	20930
	ATCGGCTGCT	GGATGTGCA	GGTCTGGCT	GAGGTCGCA	GGGTCAGGCT	TGAGTGGAT	GGGTCAGGCT	21000
	CGAGTGTGCT	AGGTCAGGCT	TGAGTGGCT	AGGTCAGGCT	TGAGTGGCT	AGGTCAGGCT	AGGTCAGGCT	21070
	CAGGTCAGGCT	CAGGTCAGGCT	CAGGTCAGGCT	CAGGTCAGGCT	CAGGTCAGGCT	CAGGTCAGGCT	CAGGTCAGGCT	21140
	CGGTCAGGCT	CGGTCAGGCT	CGGTCAGGCT	CGGTCAGGCT	CGGTCAGGCT	CGGTCAGGCT	CGGTCAGGCT	21210
	CGGTCAGGCT	CGGTCAGGCT	CGGTCAGGCT	CGGTCAGGCT	CGGTCAGGCT	CGGTCAGGCT	CGGTCAGGCT	21280
75	TGCGGTGAGCT	TGCGGTGAGCT	GTGCTGGAT	GTGCTGGAT	GTGCTGGAT	GTGCTGGAT	GTGCTGGAT	21350
	TGATGCTGCT	GTGCTGGAT	GTGCTGGAT	GTGCTGGAT	GTGCTGGAT	GTGCTGGAT	GTGCTGGAT	21420
	GATGTGCTCT	GTGCTGGAT	GTGCTGGAT	GTGCTGGAT	GTGCTGGAT	GTGCTGGAT	GTGCTGGAT	21490

00582246.062100

	GTCCGGATGG	TGCAGTCCG	GGGTGAGGTC	ACCAGGCCCT	CGGTGATCTG	GATGTGGCAT	GTCTTCTCG	21560
	TTTAAGGGGT	TGGCTGTGTT	CCGGCCGAG	AGCAGCGCTT	GGGTGAGGAG	ATCTCGGCCA	AGTTCTCTGA	21630
	TGGCTGATG	AGCTGTATCG	TCTCTGAGCT	GCTCAGGCTT	TCTCTTTATG	TCACGGAGAC	CAGCTTTCAC	21700
	AAAGAACGCG	TCTTTTCTCA	CCGGAGAGAT	GTCTGAGGAC	AGTTGCAAG	CATTGGAATC	AGGTACTGTA	21730
5	TCCCCACGCG	AGGCTCTGCG	TTCCTGAAGT	CCTGGAACAC	CAGCCCGGCC	CAGCATGTCG	CTGTCTCCA	21840
	CTTCCGCTGG	TCTCCGCGCG	TGTGACGCTC	TGGGCTGGGA	CAGCGGGGCC	CCGCTCACAG	CGCTGCTGCA	21910
	GTGATGTTCT	TGACAGGCTC	TGACTGCAGT	GAGCTCACGT	TCTCTTACTT	TGAAATACAG	GAGTTTGTGC	21980
	CAAGTGTGCT	TCAGAGGCTT	TAAAGCAGAA	GGGATTTRAA	TGATGTGAA	ACACTACCCAC	TAGGCTCTCT	22050
	GCTTTTGCTT	GGATGTATCG	TCTCTGATCTG	TCTTTTCTTT	TTTGAGATAT	AGTCTGAGAT	AGCTTCTGTA	22120
10	TTTGCCCGAC	CTGGAGAGTG	AGTGGCATAA	TCTTGSGCTA	CTGCAACCTC	CACTCTCTGCT	GTTTAAGGCA	22190
	TTCCACAGCC	TGACGCTCCT	AAGTAGCTGG	GATTACAGCG	ACCTTGCCAC	ACGCTTGCTT	AATTTTGTGA	22260
	TTCTTAGGAG	AGACGGGGTT	TCACATGTTT	GGCCAGGCTG	GTCTCGAATC	CATGACCTCA	GGTGATCCCA	22330
	CCACCTTGGC	TCCCAAAGT	GCTGGGTTTA	CAGGCTAAGC	ACCCGTGACC	AGCCCCGAT	TCTCTTTTAA	22400
15	TTCATGCTGT	TCTGTATGAA	TCTTCAATCT	ATTGGAATTA	GGTCATGAGA	GGATAAAATC	CCACCCACTT	22470
	GGCGATCTAC	TGCAGGGAGC	ACCTGTCCAG	GGAGCACCTG	GGGATAGGAG	AGTTCCACCA	TGAGCTACCT	22540
	TCTAGTGGCG	TGCATTTGAA	TGGCTGTGAG	ATTTTTCTCT	CAATTTTGCG	CTGATGAGAG	TGTGAGATTG	22610
	TGCAGCAATC	AAGCTGGATT	TGCATCAGTG	AGGACGGGGA	CGCTGTGTTT	GGGAGATGCT	AGCCTGGGAT	22680
	AGCCCGAGCC	ATGGATTATG	CTTCTCCGTC	TCCCGCCGAG	GCTGACTGTC	GAGGGGTTTA	TGTAGAGAAAT	22750
20	CAGGGCTTCC	CCAGCTCCCC	TGCACACTCG	AGTCCTCTGG	GGGCTTTGTG	ACACCCCATG	CCCCAAATCA	22820
	GGATGCTGCG	AGAGGGAGCT	GSCAGCAGAC	CTCTGTCAGG	GTAAACACAG	CTTGGGCTTG	GGGACCCCGA	22890
	CTGTGCTGTC	GGCGCATTC	CTTGCATCTG	GGGAGGGCTC	AGGCTTTTCC	CTGTGGGACG	AGTTTAATAC	22960
	ACAAATGAGC	TTACTTAGAC	TTTACACGTA	TTTAATGTGT	TGCAACCCAA	CATGCTCTCA	TGACCAAGAT	23030
25	TTTGGAAAGA	ATTTAATTGG	GGTGACCCGA	AGGACGACAG	AGACGTGGTG	GTCCCCAAGA	TGCTCTTTGT	23100
	CACATCTGGG	ACTGTGTGTC	TGCTTGGGGG	GCCTTGGAGG	CCCTCTCTCC	CTGAGGAGAG	TACGCTGGCT	23170
	TTTCTACTCT	GCTGGGCTCG	CGGCTTGGCG	TCAAGGCCAC	CGCTCCGAGC	CAGCCGCGGC	CCCAATGTCC	23240
	ACGGAGTGCC	AGGCTGTCCG	CCACAGATGC	CCAGGTCCAG	GTGTGGCCGC	TCACGCCCCC	GTGCCCCCAT	23310
	GGGTGATTTT	GCTGGATGTA	GGTGTGACGA	GAGGATGTGT	GGCTCATGAG	GGCTCATGAG	GTGCTGCTG	23380
	CTGCTTGTCG	TGAGCTCGCC	TGAGCAGCCT	CTTCCGCCCT	CTCATCATGA	AGGATGTGTG	GTCTTTCTAC	23450
30	CTGGGGGCTG	CTGCTGGGCG	CAGCCTTGGG	CTACCCCAST	GGCTGTACCA	GAGGGACAGG	CATCTGTGTA	23520
	GGAAGGGCAT	TGGTCTACGT	GGCCCCAGAT	AGCAGGCTGC	CTGGTGCTGT	CTGGTGCTGT	TGGTGGTGCA	23590
	GTCCACCTGG	GGGTGTACCG	CCGAGTCTGG	GCTGCCAGG	GTTGACTATA	GGCCCAGGTT	TCCAGGTGGT	23660
	CTGCAAGTAG	AGGGGCTCTG	AGGGGCTCTC	GGCTGGCATG	GTGTGACGTG	GGCCCGGGCA	TGGCTCTTAC	23730
	CTGTGTGCTG	CTGTGGTCTG	CGACTGTGCT	CTGTGATGTC	TGGGGGCTTG	TGGGCTTCGG	TGAGCTCTCT	23800
35	CTTACTCTCT	TCTCTGCTCG	AGCAAGCCTC	CTGAGGGGCT	CTTATTATCA	GACACACGTT	GAAGAGGCTG	23870
	CAGCTGGGGG	AGCTGTCCGA	AGCAGAGGTC	AGGCAAGCAT	GGGAGGCCAG	GGCCGCCCTG	GTGAGCTCCA	23940
	GACTCCGCTT	ACTCCCCAAG	CCTGACGGGC	TGCGGCCGAT	TGTGAACATG	GACTACGTGT	TGGGAGCCAC	24010
	AAGCTTCCCG	AGAGAAAGA	GGGTGGCTGT	CTTTTGTTT	AACTTCTCTT	TTAAACAGAA	GTTGGTTTGA	24080
40	GGCCCCATTT	TGGTATCAGC	TTAGATTGA	GGCCCGGAGG	AGGGGCCGAC	GGACACACAG	AGGGCCATGG	24150
	CAGCGGCCA	ACCCATTGTT	GGGACAGGTC	AGGTGGCCGA	GTGCGCGGTG	CCTGCAGAAA	AGAGAGGTGG	24220
	GGTGTAGGG	GGAGCTCTCG	GGGACAGGAC	AGGCTCTGAG	GACCACAAGA	AGACGCCGGC	GACAGGCGTC	24290
	GTGCGACAC	GGCCCGAGGT	CTGTGATCCG	TGTCTGCTGT	TGGTGCCGAG	CTTCCGCTGG	TTTCCGCTTA	24360
	CAGGCGCCGG	GGACACAGCC	ACGACTGCCA	GGAGCCCACT	GGGCTCTGAG	GATCTGGGAC	CTTCCGCCAC	24430
	GGCTCTTGCA	ACCCACCCCT	TGGGCTGGGG	TGGCTGCGGT	GACCCCTCTA	GTGTGAGAGA	GTGTGGGGTG	24500
45	AGCTGACAG	AGCTGTGCGA	TGAGATTCCC	GTGTGTCACG	CACATGCGGC	CAGGAGAGCC	TTTCAAGAC	24570
	GGTGTAGAGA	AGCTGGAGAG	GGTTCTAGGT	CCGCGGCTCG	GGTGGCTGGG	GACACTGTGT	AGGGGCTGCT	24640
	TCTCCCTGTC	GGCTCATATG	TGGGGTGGGG	ACTTGGCCGG	ATGCATTTTC	GCTACTGTCT	CCCATGCTGT	24710
	CCCGCCGAGG	CCGAGCTGCT	CACCTCGAGG	GTGAAGGACG	TGTACAGCTG	GCTCAACTAC	GAGCGGGCGC	24780
	GGCGCCCCGG	CCCTCTGGGG	GCCTCTGTGG	TGGGCGTTCG	GATATATCAC	AGGGCGTGGG	GCACCTCTGT	24850
	GTGCTGTGTC	CGGGCGCCAG	AGCCGCCGCC	TGAGCTGTGAT	TTTGTTCAGG	GGGTTGCTGG	GAGCCCTGGT	24920
50	AGTACAGCT	CTGGAGACCT	GGCAATGGCT	GGCTGATGAG	CACCTCTGTT	TGGGTGAGAG	AGGTATCTCT	24990
	GGTGGCCGCG	CAGGAGTGTC	AGGTACCCCT	GTCACTGTGT	AGGACACACC	TGGCACCCTG	GGTGGAGGCG	25060
	TTACGCTGTC	CTGACAGAC	ATGGGCGCGA	CTGTGCACGC	TGACTGGCCG	GTCTCTTATT	CCCAAGAGGG	25130
	GTCCCACTGG	ATTCCAGTTT	CCGTCCAGAA	AGGAACCGCA	AGCGCTCAGC	CACAGCGGCC	CGGTGCTCTG	25200
55	CACCCAGTCT	AGTCCAGAG	GGTCTCTGTT	CTGAGGCTGC	AGAGAGGGGA	AGGACGCCGC	CTGCCCCCTT	25270
	GGGTCTGGAG	TGGTGGGGGT	CAGAGAGAGA	TGGTGGGAGA	CCGCGACGCC	ATGACCTGGG	GGAGCAGGTC	25340
	ATGTCTGAGT	CTGCTGCTGC	CACATGTCAG	CTGCTCTGCA	CTGCTCTGCA	CAGGTCGATG	TGACGGGCGC	25410
	GTACAGACAG	ATCCCCACAG	ACAGGCTCAC	GGAGGTGCATC	CCGACATCCA	TCAACACCCA	GAACACGTAC	25480
	TGGTGGCGTC	GGTATGCCGT	GGTCCAGAA	GGCGGCCATG	GGCACGTCCG	CAGGCGCTTC	AAGAGGCCAG	25550
60	TAAGTTCAC	TAGTTAATG	CTGTCCAGG	ATGTGTGTTT	TGGGGTATG	ATATGTGATG	TAAGTACGAT	25620
	GTGTGCTGTA	TGGGTTCTTG	TGGTGGAGTT	ACTTCCAATG	TTTACACATC	TGTGATATCT	GTGTGTGGCA	25690
	CGTGTGCTGC	TGGTGGCATG	TATCTGTGGG	CTGCTCAATT	TTTGAGTGGT	TGTGTGTGGC	GTGTGTGTGT	25760
	CCATGGGTGG	CTGCTGCTGC	CTGCTGCTGC	GTCTGCTGCT	GTCTGCTGCT	GTGATGATCT	CTGTGTGCTG	25830
	CATGTCTGTC	AGTGGCTATT	TGTTGCTGTT	TGTTGCTGAT	TGTCTGTGAC	ATATGCGTGT	CTATGCGATG	25900
65	GTGTTGTGTC	GCCTCTTGGG	CTTACTCTCT	CTCTCTCTGG	GCATGTTCCG	CACCATTTGT	CTACAGGCTCT	25970
	CGGGTCTGGG	TCTGGGAGC	TCCCAATCA	GGGTCTCTAC	TCTTAGCATG	GGTGGCCCTG	TCTGTCTACA	26040
	GGGCTGGGCG	TGGAGACCTC	TAAAGCAGGT	TGTAGAGAGG	ATGTAGGATG	CTTGTGTGAT	CTTCTGGAC	26110
	CCCTGGGACG	CCGAGACACC	CAGTCTGGCC	TAGAGCCGGT	GTCTGCTGCT	CATAGAGATA	TAGAGAGGCT	26180
	TGCGTCCGCG	GGACACACTC	CTCCACAGAC	GGCGCGGGTG	CTTGGGGCTG	GGCAGAGGCT	AAAGGGGCCC	26250
	TGGGTTTGGG	TCCCAACCCA	GTGTCATG	GCACGCTGGA	GGGGTAAACC	CTCAAGGTGC	TGCGAGCGGC	26320
70	GGGTGACAGT	GTGAAGAAGT	ATGCTCTGAC	CTTCCGCTG	GGGAGAGGCA	CATGTGGAAA	CCACACAGGA	26390
	CTCTCTTCTG	TGACTCTCTG	AGCT					26460

Contig 2:

	TGTGGGATTG	TTTTTCATGT	GTGGGATAGG	TGGGGATCTG	TGGGATTGGT	TTTTATGAT	GGGGTAACAC	70
	AGAGTTCAGG	CGGAGCTTTC	TTCCCTGTAGT	GGGTCTGCGG	GTGCTCCACG	AGCTTTATIT	AGSAGACAT	140
5	ATCTTCCTTT	CACTCTGGGG	GGGCTTATTA	GTGTGTGATG	GTGTGTGGGG	GTCCTCCCTG	GCTCCCTGTT	210
	CTGTTCTCTC	CTCTCTGGGG	TGCTGTGGTG	CCCTCTGCGG	TGCTGTGGGG	GTGGGCGAGG	CTTCCGCGG	280
	TCCTTTGTGT	CAITGGGCGT	GATGTGGGCC	TGGCTACGCT	CCGCTCTTGG	AATTCGCCCT	CGAGTGTGAG	350
	GCTTTCTTTC	TTTCTTCTTT	TTTTTTTTTT	TGATAACAGA	GTCCTGCTCT	TTTTTGGCCA	TTTTTGGCCA	420
10	GGCTGGGATG	GTTTTGGCGT	ATCTTGGCTC	ACTGCAACCT	GTGCTTCCCT	AGTTCAAGCA	ATTTCTCTGC	490
	CGCTACGCTC	GAATATAGG	GAATATAGG	GCGCCACACC	CATGTGACT	AATTTTTGTA	ATTTTAGTAG	560
	AGACGGAGTT	TCCTCATGTT	GGCCAGAGGTS	GTCTCGAAGT	CTTGACCTCA	GGGTGATCTC	CACCTCTGGC	630
	CTCCCAAGT	CGTCGGATGA	CAGGTGTGAA	CCGCTCTGCC	CGGCGAGAC	TGCTCTCTCT	CAGCTTCGCT	700
	GAGATCTGCA	GGCATAGCTC	CTGCGACGCT	TGCTGCTGAC	AACCTCCGTT	TTCCTCTCTC	AGGCTCTGCT	770
	AGGGGCTCTT	CCATTTTCAT	ACTCTCTTCA	CAGAGAAGTT	TCAGCTGTGC	TGATTTCCCG	GCTGTTCCTC	840
15	GGCTAATTTG	TGTCCTGCTT	TTATCAGTGG	CCTCCTTCCA	TTTCTTTAG	GCTTTGTTTA	TGTTTGTGTT	910
	TCCGGCTCTT	TGAAGAAAAA	GTTCCTGATTA	TGAGTGTGTT	AACCTTCTTT	TCTAAACACG	CAITCTGAAT	980
	CTCTCTTTTC	AGGATACCGG	AGGATACCGG	AGGGCTGTGG	GTCTGGAGTG	GCACCGGTCT	GGGCGCTGTT	1050
	AGGAACCCGG	CGCAGACGCG	GAGGCTAGGT	GGGCTGTGGG	GAGTACAGCT	TCTCCTGTGA	GGCCGCCGCC	1120
	TCTCAGATCA	CGACCTGGAT	GGGCTGTGCTA	GAGGCGCACA	CACCTTACTG	AGAACCTGTG	GTGAGGGGCG	1190
20	TCTAGATTCT	GTGCTCTTTA	TGGGAATCTA	ATGCCGTATG	ATCTGAGGTG	GAACTGTGTT	TCTCCAAAAC	1260
	CATCCCTCTC	CCCACTGCTG	TCCCTGTGGA	AAATCGTCTT	CCACGAAACC	AGTCCCTGGT	ACCACAAATG	1330
	TGTGGGACCG	TGTGCTTAAG	ACCTGCTTCA	CGAGCTCTCT	TCAGTGTGTT	ATATATTTGG	TATTTGTGTT	1400
	TGAGTCCAGA	ATAATTCGGT	ATTCTTGTGA	GTCTTTCGCG	CGACTCTAGA	CCCATGGGCT	ATTTGTGGGG	1470
25	CTCTTCTCTG	CTCTCGGGTT	GGGAAGGGTG	CAGAGCCCAT	GTACTTCTCT	GTTACTTCTCT	TCCAGTGTGG	1540
	TTCTCAGGGT	TGAATCTGAC	TGGATGTGGT	TTTAGCCCAAC	GGCCTTGGCC	CCAGTCTCTG	GGGCGTGGGG	1610
	AACATCTGTA	AGCAACAGAT	CACCTGTGCG	GTCTTTTGAAT	GCCTCACAAG	CTGAGGAGCT	CCCTGTGTCCG	1680
	TGTTAGTGTG	TGCTCAGCTG	TGCTCAGCTG	CTGCTCTTGG	GGAGCGAGGG	CTCTAGGACG	TCCCGTAGTA	1750
	ATATGACAGC	GTCTCTGGGG	AGTCTGTGGA	ATAGAGAGGTT	GGGGTCCCGG	TCTCTTCCCC	GCCTGTTCAG	1820
	CTCTTCTCTG	CACTCTGATG	GTGCTGTGCA	TGAGATGCTG	CCATATCCCT	CAGACTGTGG	GTGCTGTGGA	1890
	CCGGGAGCTC	GAGTGGCACT	GTGCCAGGCT	GACTGTGGAT	CCAGCGGGCT	TGATGTGTGG	TGATGTGTGG	1960
30	TGACTGTGGA	TGGCGGTTGG	TCACAGGGGT	CTGATGTGTG	GTGACTGTGG	ATGGCGGTGG	TGGGGTGTGA	2030
	TSTGGTGTGT	TGGTGTGGGG	GTCTGGGGGT	CTGATGTGTG	GTGACTGTGG	ATGGCGGTGG	TGGGGTGTGA	2100
	TGTGGTGAAT	GTGGATGGCG	GTCTGGGGGT	CTGATGTGTG	GACTGTGGAT	GGCGGTCTGT	GGTGTGTGAT	2170
	TGTGAGCTGT	TGGTGGAGCT	CTGGGGGCTT	GATGTGTGGT	GACTGTGGAT	GGCGGTCTGT	GGGTTGTGAT	2240
35	TGTGAGCTGT	TGGTGGAGCT	GTCTGGGGGT	CTGATGTGTG	GTGACTGTGG	GATGCGAGCT	TGGGGTGTGG	2310
	TSTGTGACTC	GTGGATGGCG	GTCTGGGGGT	CTGATGTGTG	GTGACTGTGG	ATGGCGGTGG	TGGGGTGTGG	2380
	TGTGTGGTGA	TGGTGGATGG	GGGTCTGGGG	GTCTGTGGGG	GTGACTGTGG	ATGGCGGTGG	TGGGGTGTGG	2450
	TGTGTGGTGA	TGGTGGATGG	TGATCGGTCA	CAGGGGCTGG	ATGTTGTGGT	ACTGTGGAGT	GGCGTGTGGG	2520
	GGTCTGATGT	TGGTGGATGG	TGATCGGTGA	TGGGTGTGTA	GGGCTGTGAT	TGTGGTGTGG	GTGGATGTGG	2590
40	GTCTGGGGGT	CTGATGTGTG	GTGACTGTGG	ATGGCGGTGG	TGTGGTGTGG	TGTGGTGTGG	GTGACTGTGG	2660
	GATGCGAGCT	GGTCTGATGG	GTCTGTGGTG	TGTGACTGTG	GATGCGGTGG	CTGTGGGTGT	GATGTGTGGT	2730
	CAGCTGGGAT	GGGCGTGTGG	GGGTCTGTGG	TGTGTGACT	GTGATGTGGC	CTGTGGGTGG	GTGATGTGGT	2800
	GACTGTGGAT	GGGCGTGTGG	GGGTCTGTGG	TGTGTGACT	GGATGGCGGT	CTGTGGGTGG	GATGTGTGGT	2870
	GACTGTGGAT	GGGCGTGTGG	GGGTCTGTGG	TGTGTGACT	TGACTGTGGG	TGGCGTGTCT	GGGTTGTGAT	2940
45	TGTGGTGAAT	TGGATGGGAG	TGCTGGGGGT	TGATGTGGGG	TGACTGTGGA	TGGCGGTCTG	GGGTTGTGAT	3010
	GTGTTGTGAG	ACTGTGGGAT	GGGCTCTGAG	GGTCTGTGAT	GGTCTGTGAT	GGTCTGTGAT	GGTCTGTGAT	3080
	ATGCTGTGGT	ACTGTGGGAT	GGGCTCTGAG	GGTCTGTGAT	GGTCTGTGAT	GGTCTGTGAT	GGTCTGTGAT	3150
	ATGCTGTGGT	ACTGTGGGAT	GGTCTGTGAT	ACAGGGGCTT	GATGTGTGGT	GACTGTGGAT	GGGCGTGTGG	3220
	GGGCTGTGGT	ACTGTGGGAT	GGTCTGTGAT	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGGCGTGTGG	3290
50	GGGCTGTGGT	GTGTTGTGAT	GTGGATGGCG	GTGCTGGGGT	GTGCTGGGGT	GTGCTGGGGT	GTGCTGGGGT	3360
	GTGCTGGGGT	GTGTTGTGAT	GTGGATGGCG	GTGCTGGGGT	GTGCTGGGGT	GTGCTGGGGT	GTGCTGGGGT	3430
	GTGCTGGGGT	GTGTTGTGAT	GTGGATGGCG	GTGCTGGGGT	GTGCTGGGGT	GTGCTGGGGT	GTGCTGGGGT	3500
	GTGCTGGGGT	GTGTTGTGAT	GTGGATGGCG	GTGCTGGGGT	GTGCTGGGGT	GTGCTGGGGT	GTGCTGGGGT	3570
	GTGCTGGGGT	GTGTTGTGAT	GTGGATGGCG	GTGCTGGGGT	GTGCTGGGGT	GTGCTGGGGT	GTGCTGGGGT	3640
55	GAAGCTTCCC	AGGCGCTCTC	TGGGCTTCAT	CCGCGCATCG	GGCTTGGCCG	CAGGTCACCA	CGCTGTGATG	3710
	GGAGAAACAA	AGTGGCCACG	TGTCGGCGGG	GGAGCGCACA	TTTTTGGGCT	AGCTTGGGCT	CTCTGCGCGG	3780
	AGGTCTCTAC	CTTGACAGAG	CTTGACAGAG	ATCATGACGA	GTGCTGTGCT	CAGCTTGGGCT	AGAGCCGCGC	3850
	CTTGACAGAG	CTTGACAGAG	CTTGACAGAG	ATCATGACGA	GTGCTGTGCT	CAGCTTGGGCT	AGAGCCGCGC	3920
	GTGGGCTCTC	CCCTGGGGCA	TGACTGGGCT	CATGACGCGA	CAGACTGTGT	CCCTTGGGCG	GAGTGTGGGG	3990
	GATAGAGCTG	TGATGGGGGG	ATGATGAGCT	GTGTGCCCTG	GGCAATCTG	AGCTTGGGCCA	TGGCAGGCTG	4060
60	GCACAGCTGT	TGATTTACAG	CACCTGTCTA	GTCTTGAATG	CGGCGGCTCT	CTCCAGTTCT	GCACTGCTCT	4130
	TGTTCATGAT	TGTTTAAATG	TTCTTCTCTG	CAGTITTTGAT	CTTGAGGCGA	AAGGAAAGGT	CCGCTCCCTC	4200
	TTTGAAGAGT	CAGGCGCTGT	TTTGAAGAGT	CTCTTCTCTG	CTCTTCTCTG	CTCTTCTCTG	CTCTTCTCTG	4270
	AGGAACCTGT	CCCTCTTCTG	TGAGGACAGG	TGAGGACAGG	TGAGGACAGG	TGAGGACAGG	TGAGGACAGG	4340
	TGCTGGGCTG	GTTCACCTGT	CCCTTGGGCC	CTTTGACAGT	GTGCTGTGCT	CAGTGTGGCT	TGTGGTGTGT	4410
65	TGCTAGAGCT	TGTTAGACAT	TGCTGGGCTC	TAGGGGAGAC	TCGTGTCCAC	CGCATGAGAG	TACAGAGACT	4480
	CTGGGCGGAA	TTCTTTGGCT	CCGAGGGTGG	GGGTGTGAGT	GGCTTGGGCT	CTGTGGGACG	AGAGCTTGTG	4550
	CCGAGGAGCT	GGGAGACAC	CTCTGGATCA	CATATGACCA	CCGCGGCGAG	GTGGGCTGTG	TGTGGTGTGG	4620
	CCGAGGAGCT	GGGAGACAC	CTCTGGATCA	CATATGACCA	CCGCGGCGAG	GTGGGCTGTG	TGTGGTGTGG	4690
70	CACTTACCCG	AAAAGGAGCG	GAGGGTCTTG	GCACAGTGTG	CTCTGCTGTC	TAGACACACA	CCGCTCTACT	4760
	CCCATGTGTC	TCCCTGTCTG	TTTGTGAGAG	CTCTTCCCTG	AATGAGGCGA	GAGTGTGGCT	CTTGTGAGCT	4830
	TCTCTACAGT	TGATTTGGCA	CACGCGCTGT	CGCATCAGGG	CGAAGTGTGT	GAGTGTGGCT	GTGTCAGTGT	4900
	GGCTTGGGCT	GGGTTGGGCG	GGTGTGAGGG	CTTCTGTCTG	CTCTTCTCTG	CCCTCTTCTC	CAGTGTGGCT	5040
	CTCTTGGGCT	GGGTTGGGCG	GGTGTGAGGG	CTTCTGTCTG	CTCTTCTCTG	CCCTCTTCTC	CAGTGTGGCT	5110
75	GGGCGCGGAA	AGATGGCTCT	CGTGGGCGCG	CGCGAGAGCG	ACGAGTGTGG	CGGAGAGGCT	TGGAATAGCT	5180
	AGGCGGGGTT	TGGAGTTGCT	CTCTGTGGGA	GGACAGAGGG	CGGCGGGTGT	GTCTGTGGTA	GGTGTGGGCT	5250

00562246.091200

	GAGCGTTTGA	GCGTCGACGT	TGTGAGCTCC	AAGTTACTAC	TGACGCTGGA	CACCGGCTC	TCACAGCGTT	5320
	GTATCTCTCT	TCGCGGATAC	AAAGAGGATT	TATCGGATTC	TCATTTCTGT	CCCTCTGCTG	TGACCCCGCG	5390
	GAGCGCGCG	GCGGCTCTCT	CTGTGACTAG	ATTTGCTATC	TGGAGGATCT	GGGGTGTACC	GTCTAGTTTG	5460
	CTCTCTCGG	GCGGCGCTCG	GTGCGCATGG	CTGCGAGCG	CTGCGAGCG	TGCGCTGACA	CAGCAGCTCG	5530
5	GTGAGGCACA	CTCGCGGGTG	TAGAGGCGACA	GTGCGTGGTG	CCACATCAGC	CTCTCTGGAT	TTTAAAGTAA	5600
	ACCACACACC	TCCCGCGACG	CATCTCGCTG	CGACCTCTGT	TGTGCTGGGG	GAGAGTGTGA	GCGAGCGAGCA	5670
	AATTCGTGCA	CTCATCAAGT	CTCATCGCA	GTGAGTGGCA	GTGAGTGGCA	CGCTCGAGGC	CTCTCTCTGG	5740
	CTGCTCTCC	AGCGGATAAA	GGAGCTTGCA	CAGCTTTCGA	AGCTTTTAT	TAAATATATAT	ACTATTTAT	5810
10	ATTGCAATAT	AAGTATATAC	AGCTATATAC	ATATATATAT	ATATATATAT	AGATATATAT	AGAAATATTA	5880
	AGTACTACAT	AGCTCTCGA	AAACACAA	TTGCGACTAT	CAGCAGATCT	ATTTTGTGCG	CTCTCTCTCT	5950
	TGTGACAATG	TGTGTAAAGC	GCGCCGAGGC	CCACAGAAAT	GCGTGACAAA	GTACCTCTCC	CAGAGAGCC	6020
	ACGACCGCGC	TCCTCTGGTG	TCGTGAATTT	TATTAAGATG	GATCAAGTCA	CGTACCGCTC	ACGTGTGGCA	6090
	GGGCTTTGGG	GAATGTAGAG	TGATGACTGC	GTCTCTATCG	CTCGACAGAC	AGGAGGTGGT	TGTGCTCTGC	6160
	CTGTCCTTAC	GACACGGACA	GCGCCGAAAG	TCATAGTCCC	ATGCTGGTCC	AGTTTGGGCT	CTGATATAAA	6230
15	ACGCTCTCAA	AACCAATCT	CCCAAAATCT	AAGACAGAG	AGAGTTCTCC	ATCCCATGTG	CTCAGCGGG	6300
	CGTATCTGCT	TGCTTGTGAT	CGCTGGGCTG	GCGGAGCTCT	TAGAGTTGGT	CGCTCTGCTT	CTGTGCGAAA	6370
	AGCGAGCTCG	CTTCTGCCAT	CACCTGTGATA	TCGACACAGC	CAAGGAAGGC	CTCTTTCTTT	TTCTTTCTTT	6440
	TTTTTTTTTT	GAGACGGAAC	GTCACGTGTT	TCGTGCTGGG	CTGTAGTGCA	GTGGCCGGAT	CTCAACTAC	6510
20	TGCAACCTCC	GCCTCCCGGG	TCACGACTT	TCCTCTGCT	CAGGCTCCCG	ACGACCTGAT	ATTACAGCCA	6580
	CCGCCCTCT	GCGGCTGGCT	AATTTTGTGA	TTTTTAGTAG	AGAGGGGTTT	TGGCATGTTT	GCGCAGGCTG	6650
	CTGTGACAT	CTGTGACTCA	GGTATGCTAC	CACTCCAACT	CTGCCAAGCT	GTCTGGATTA	GGCTGTGAG	6720
	CCATACCGCC	CAGCGGGAAA	GCCTCTTTT	AAGTGAAGA	CTATATGCG	TTCCCGAAA	TATACAGCT	6790
	TGTTTTTGA	GTAGGCTGCA	AGCGTCTCT	AGCAACGACA	GTGGCTCTCT	GTGGGCTCTG	GGAGTGTGCT	6860
25	AGGGTGGGCT	GGACAGCAT	CCCTCTGTGT	GCACCTTTAG	GTCTCACGGG	GCTATTTCTG	TCTCATGTT	6930
	TGCTGACAAA	GCAGCCCTTG	GCATCCTGT	TGGAGAGTTT	TCTGCTTCT	TGTGGTCAIT	GTGAACATG	7000
	GGGCAAGGTT	TATTCGCTGT	GCGCGCAGCG	GCTACATGTA	GGGTCATGAG	CTTTTTCACG	TGACAAATTT	7070
	CTCTGAAAA	AAATGAGGA	GCTCTACAT	GCTCTACAT	GGGTCACAT	CTCTCTTCTT	CTCTCTTCTT	7140
	TCTGAGATT	AGAAATCTT	AATGAAGAAA	AACTCTTGAT	ATTGACAGCA	AGGATGTGGT	CCACCTCTGT	7210
	CTGTGATCTG	TTTCACGCGC	CCGACGTGAT	GGTGAGAGTG	GGAGCAGGCG	ATTGTTTGTG	CAGAGTGTCT	7280
30	ATCTGGATATG	CTCTGAGGTT	GTTCGGCGGG	TGAATGGTAG	ACGTGTCTGT	TGTGATCTAT	AGGTTTCTGT	7350
	TCCTGTGTTG	CTCGGTTGTT	AGTGTACGCA	TCCTCAGCAC	ATCGCTCTGT	CGTCTCTCAC	CTGTGCTCTT	7420
	CCGCGCCGAG	TCTGATGCTC	AGTGCAGAGG	GATCCGCGAG	GGCTTCATCC	TCTGCAAGCT	CTCTGTCAGG	7490
	CTGTCTACAG	AGCAGATGGA	GATCTGAGAG	TTGCTGAGAG	CGTGTGAGCG	CGTGTGAGCG	CTCTCTCTCT	7560
35	CCAGSGGCGT	CTGCTGGGTT	GTGATTTGCG	TTTTGATCTG	TTACAGTGTA	ATATCTCTGG	TCCTCTGGAG	7630
	ACCATGATCT	TCTGCTCTGT	AGGAACACGA	CAAGGTTGCA	GCGCTTCTTA	GTATGACGAG	CGCAGCGGAC	7700
	GGGTGTGACA	CTGTGAGGAC	TGCGGGGCTAC	ACCGCAGGCT	TGTCCAGCGG	CCATGTCCAG	AGGCTCTAGG	7770
	CTGACGAGCG	CGGAGGGGCT	GCTGCGCTGC	ATGATGAGCA	TGTGAATTTCA	AGCAGCAGGA	AGCACACGAC	7840
	CTTTCTGTCAT	CTGACCAAGG	CTGCTTTAGG	GTATCTTGAG	AGGATGGGCT	GTCGCAAGCT	CTGCGAGCTC	7910
40	ATCTCCGCGT	AGCGGCTCTG	CAGCTGCTCT	CAGCTGCTCT	CTCTCTGAGC	CCATCTGCGA	CTGCACTCTG	7980
	ATGSGGCTCT	CACCCAGACA	CGCACACACC	TAATATCTGT	GCCAACTTAA	TGTGCTTCAA	CTCAGCTGCG	8050
	TTTTTGTGAC	AGCAGATTA	TTTTTTTTTT	TAATATCTTA	AGTCTTAGGG	TACATGTGCA	CGAGCTGAGC	8120
	GTATGATGTA	TATGATATGA	TGTGCAATGT	TGTTGTGCTG	CACCCATTAA	CTCATCTATT	ACATATGATTA	8190
	TATCTCTCAA	TGCTATCTCT	CCGACATCCC	CCCATCCAT	GACAGCGCCT	GGTGTGTGAT	TTCTCCCACT	8260
45	CTGTGTCGAA	GTGTGCTAT	TGTTCTGATC	CCACCTGTGA	GTGAGAACAT	GTGTTTGTG	GTTTTCTGTT	8330
	CTTGCANTAG	TTTTGCTCAG	GTGATGGTTT	CCAGCTTCTT	CTCATCTCTC	ACAAAGAGCA	TGACATCTATC	8400
	CTTTTTTATG	ACTGCTATAT	ATTCCGGGCT	GTATATGTGC	CACATTTCT	TAACTCAGCT	TATCATCTAT	8470
	GGACATTTGG	GTGGTTTGA	AGTCTTTGCT	ACTGTGAATA	GTGCCGCAAT	AAACATGAGT	GTGCTATGCT	8540
	CTTTATAGCA	CGATGATTTA	GAATCTTGT	GGTATATACC	CAGTATAGCT	ATGCTGGGCT	CAATGTGAT	8610
50	TTTTGATGTT	AGATCTCTGA	GGAAATCACCA	CAGTGTCTCT	CAGATATGGT	GATGCTTACT	AGCTCTCCAC	8680
	CAACGCTGTA	AAAGCTCTCT	CTGCTGTGAG	ACGATCTGTA	CAGATCTGTA	TTTTTATGTA	GAATGATATC	8750
	ACTGACACAG	CAGACAGCTA	GTGAGAGGAT	CGTCAGAGAG	CTGACAGGCC	ACACAGCCAT	TTCTCTCTCA	8820
	GAGCTCCGGGT	TTTTCTCTGT	CATCTTTTGA	AACTCTAGCT	CCAAATATAG	CATGTACAGT	GGATCAAGGT	8890
	TCTCTCTCAT	TAGGTTTCAA	GTCTAGATTT	GAATTAAGTT	TATGTAAACG	AAACAAATAT	TTTTCTGTGA	8960
55	CACACATCT	TCCTGGATTT	GGAGGAAAGT	GTCTCTGAGC	TGGCGAGCA	CTGCTGAGCC	TTCTGGAGCA	9030
	GGATATCTCT	GCGGCGATGT	CATGSGGGCT	TGGGCTTGGG	CTGAGGGTTC	CAGATCTGCA	CGATGCGAG	9100
	CTTCTCTGCG	TAGAGATCT	GGATGAGTCT	CGACAGAGTG	TAATGAGGCT	CCATCTCTGCT	CGATGAGGAG	9170
	ACGCTGCTCA	GAGGGGGCGA	GGTCCGACG	CCGAGCTTCT	TACCTGCTCT	CACATTTATT	TCCTCAAGAG	9240
	TCTSAGAAT	GGGGCGCGCG	CTGATGGGCT	TCGTTCTGCT	TCAGCTGGCA	CAGAAATGCA	CGAGCTGTAT	9310
	GTAACACACT	GATGATTA	ATAATGAGG	AATTTGCTGA	CGAGTTAACT	ATGAGAGAGT	CACTGTCTGG	9380
60	AAAGAAATTT	AAGTTTCTCA	TTTACCCGCT	TGGGGAATG	TATCTTTAT	TATGCGCTGT	TAAATTTT	9450
	GACANTCAGT	CTGCTCTGTA	CATCTCTGCT	TGAATAAGT	TAAAGTTAAC	CTTCTCTGCT	ATTCTCTGCT	9520
	TTTATGAGCT	TGCTCTGCT	TGCTCTGCT	GTGAGAGCT	CTGACAGCT	CGACTCACCC	CGGCGAAGAC	9590
	CCGCTGCTG	TGAGGCGCGT	GCGCTGTCT	TGCGGAGCC	TCGACAGCCT	GTGGGCTTGT	CAGTTGAGCC	9660
65	CCGCTGTGCT	TGCGCTCTG	ACCGCAGGCT	TGCTCTCTGC	AAGTCTCTCT	CTCTGCGCG	TGCTGCGGCG	9730
	GGAAGAGCAG	AGGCGCTTGG	CGGATGAGCC	AGGCTGTGGG	GCGCAGGCGC	ACCTCTGGGA	GGGAGTGGGT	9800
	ACCGTGGAGC	CCCTGCTGCT	GCGAGAGGCT	ACCGAGGTTA	CACAATGGGT	GAGTGCAGCG	GTGACCTGCT	9870
	CTCTCTGCTG	CTCTGGAAT	GTCAAGCTG	GCGCTCTG	GCGCAGGCT	CGGCTCTGCT	GAGGCTCTG	9940
	CCAGSGGCTG	CAGGGGCGAG	GCGGCTCT	CTCTGCGAG	GTGACCTGTA	CGCTGGGAGT	ACGAGGAGCT	10010
	CTCTAGTATG	CTGCGCCACA	CGGTTGCTGT	CGGTTAGCTT	CTGTGCTGGG	AGTTTGGG	CTGCTGGGA	10080
70	GAATTTGGA	TGCTGAGTG	CTGCTGTCTT	GAATCTAGG	GATGCTAGG	AGTGGGTTCT	AGAGTGTATT	10150
	TTTTGTAAAT	AAATCAAAAT	CAGGTAAGG	GGAGCTGAGC	TGAGCAGAGG	GATGTCTCAA	AGTGTGCTCC	10220
	CTCTAAGGGC	GCGCCACAGA	CGGCTGTGGG	TGCTTTTAAA	TGCTGATTTT	ACGAGAGGCT	AGAAACCTCT	10290
	AAATCTCTTA	CAGGCTGAG	CGGCTGTGGG	CTCTGCTGAG	GTGACCTGTA	GGGCTCTGCT	CGGCTCTGCT	10360
	TTGTGAAGAC	CAATTTGAAAC	GCGGCTCTCA	ACTGCTGAGC	CAGTCTCAC	ACGAGGAGGC	CGCCTCTGGC	10430
75	TGGGGGTATG	CTCGGCTCTT	CTCTGTGGCG	AGCCGCGAGC	AAGTAGGCT	GTGCACTATT	AAATCACTTA	10500
	AGTCTCACT	GCGGCGAGCC	CAGGTTGCTA	GCAATGAGG	GCTCAGAGCT	CTGAGGCTGT	CTGAGGAGCC	10570
	AGGACAGAGC	GCGACAGCTG	CTCTGTGCTG	CGAATGCTG	GAGGTTGCTG	CGGACAGGAG	TGGCTACAGG	10640
	TGTATGTTGG	GCTGCCACCG	GCGGCTGAGC	GCGGCTTGTG	ATGATGCGG	AGGCAATGTA	CAGGAGGGGG	10710

	TGGCCACAGG	GAGCTTGGAA	TGCACACAGG	GAGCTGCGCA	GCTGGCCGAG	GCTCCACAGG	CAGGCCACAG	10780
	GAAAGGACAG	GGAGGACCCG	GGGACACAGG	AGAGGCGCA	GGAGGCGCA	GGAGTGCCTA	GGAGTGCCTA	10850
	GAGCTTACG	GGACACAGG	GGCTCTCTGA	GCTGGGTGCT	GAGCTGCCTA	GAGCTGCCTA	GAGCTGCCTA	10920
5	TTGACGTGAA	GCTGACAGCT	GGTGTTCGCC	AGCTCACAGC	CACAGCAGCT	CCCGCCGCTG	ACGAGCAACT	10990
	CAGAACCCCT	CCCTTTTGCT	AAAGCACAGC	AGATGCCCTC	AGGACATCTA	GGAGAAACAA	GGCAAAATGC	11060
	TTGAGAAACG	TCCTTAAAGA	AGGTGGGATG	TGTGGCAATT	TTTGTCCAGA	CTTGTGCTG	CCCGCCGACA	11130
	CAGATGAGTC	TATAACAGGA	TTGTGGTGTG	GGCATGGGGA	CACATGAGAT	GGACATCAAC	AGAGGCCACT	11200
	GGGGTGTGAC	CTCCGCTGCT	AGTCCCTGCT	GTCCCGGGTC	GTCAGCAGGT	TTTGTGATCT	TCACCTACCT	11270
	GTCTCTCCCG	GGAGACAGG	AAAGCACTCC	GAGCTCTGGA	GTCCAGGCTG	GTCCAGGCTG	CTCAGAGCTG	11340
10	GTCCACAGCG	CAGCACCGTG	CTCCAAATCA	CCACTTCTCT	GGGTTTTC	AAAGCATTTA	ACAGAGGTGT	11410
	CAGGTTACCT	CTTGGGCTGAC	GGCCCCGCAT	CTCTGGGCTG	CACTTGGCTC	AGCTTGTCTT	GACCTGTGTC	11480
	GAGGCTGTCC	CTGAGTATGG	CTGCTGTGGT	AACTTGCAGA	AGACAGTGGT	GAACCTCTCT	GATAGAGAAC	11550
	AGAGCCTGGG	TGGCACAGG	TTTGTTCAGA	TGCGGGCCCA	GGGCTTATTC	CTCTGGTGGC	GCCTGTGCTG	11620
	AGATATCCCG	ACCTCTGAGG	TGCAGAGCGA	CTACTCCAGG	TGGCGGCAGC	CTCTGGCGAG	TGGAGGCTGT	11690
15	SCGCTGTGG	GGCAGTGCT	CTCAGAGG	GTCTGGCTCC	ACCTCTGTCC	CCGCTGTGGG	CAGGCGACTG	11760
	CCAATACCCG	AGSGTCAAG	GGCACAGGCT	CCGCTCTGCT	CCACTGGGG	TGCAGACAGA	ATGACTCTTT	11830
	CTGTGGGAGT	GAGGCTGCTC	ACAACGGGAG	CAGTTTCTTG	TGCTATTTCG	TGAAAGAAAG	ATGGTGCAAC	11900
	AGACCTGGGT	AGACTGAGT	GTCTTCAGAA	AGCAGTCTGG	ATCCGAACCC	AGAGCGCCCG	GGGCTGTGCT	11970
20	GGGCTGAGTC	TCTCAAAACC	GAACACAGGG	GCCTCTGTCG	GCATGAGTGC	CTCTGAACCC	GAGACCTGCT	12040
	GGGCTGTGCT	GTCTGGAGTC	TCTCCGAACT	CAGAGACTTC	AGGGGCTTTT	TGGGCGTAGT	TCTCTCCGCT	12110
	GTGAGGCCA	CATCTCAAGG	CTCATCCACA	GTCTCAGAGA	TGCATGAGT	TGATGATCAC	GTGTGACACA	12180
	TCAGGGGACG	GGGCTCATGT	GTGGGGGGGG	TCTCTACAAA	ATTCTGGGGT	CTTGTTTCCC	CAGAGGCCAG	12250
	GAGCTCAAGG	CCGCTCTCTA	GGCTCAGACA	CAAAATGAAT	AGAGATGGAC	ACAGATGAGT	AAATCTGTGC	12320
25	TGTTCTCTTT	ATGATAAAAA	AGTATCAACA	TCCAGGGAG	GGCAAGGTGG	CTCCACACTA	TAATTCACAG	12390
	AGTTTGGGAG	GGCAGGTGGG	GTGGATCAGT	TGAGGCGCAG	AGTTTGGAGC	CACACTTAAC	ACAGATAGTA	12460
	AATTCCTATT	CTACTTAAAA	ATACAAAAAA	TTAGCTGGG	GGGTGGCGAC	AGCTGTGTAG	TCCGACTTAT	12530
	GGGGGAGGCT	GGGACGAGG	ATCATTTTGA	AGCAGAGGTA	CAGAGGTGTC	AGTGGAGGCT	AGCAGGCTAG	12600
	CTCAGCTCCA	GGCTGGGCAA	CAGAGTGAGA	CTTCACTTTA	AAAAAATAAA	AAAAAATGAT	AGCATTTCAA	12670
30	AACCATGATG	GACAGGCTGT	TTTTTATTCT	GTCTTCGAT	ATAATTACT	GGTGTGTGTC	TAGAGGGCGC	12740
	AACCTGGGGT	GGCTTCTCT	GAAGGCGACA	CTTCTCAAGT	AGAGAAATA	ATGTTGTAAT	GGTTGTAA	12810
	CACAGAGTTT	AAATGGGGT	CTCTGTCTTC	TGAGTTAACA	GTCCAGATCT	GGACTTTGGC	TCTTTCACGA	12880
	ATGCTTCTGT	GGGTTTCTCT	CTTGGGGGAG	CAGCAGTGT	GTGAGCCCTC	GTGATGGGTC	AGCAGGAGT	12950
	GCAGAGGCT	TCATATGGGG	GGATGGGAC	GTGGAGAGT	CTCTGTGCTG	GTGTCGAGC	ATGTCTGTCT	13020
35	TGCACTCTCC	TCCCACACAG	GATGCCGGTC	TCTGTGTGTC	CCCAAGTCTC	CTGCTCTCTC	CTCAGACAGT	13090
	TACCTGGTCT	TGGCTTCTAC	TGGCTTTGTC	TGCATGATT	CACATTTTTC	TGGGCTCCCC	GACCTCTCTC	13160
	GCCTCTCCCA	GGCACTCTCT	CAGTCTGCTG	CATCAGGCTC	AGCTGTGAAC	TGTCCACATC	TTATTTTCTG	13230
	CCCATGAAAA	TGATTTTTTT	AGGACAGGCA	CCCTCTGGTC	CAGCCTCTGC	CAGCAGCTCA	GTGATTTGTA	13300
	TGAGAGGACA	GGTGGGAG	AAACAAATTA	GGAGCTCTGA	ACACATCTGA	ACACATCTGA	TCCGACTTAT	13370
40	GGCTTGTGCA	GGCTGGGTGT	GCATTAAGTC	ATCAGATGTC	GCTCAATGTC	CAGATATATT	TGTCTCTACA	13440
	AGGACCACTT	GGTCAAGAGT	TGTGCTTGCA	GAGGTGGCTC	TAAAGGCTCA	GCTGTGAGG	CAGTGGTTCG	13510
	CTCATCTCAG	GGTGAACCTA	CATCTCTCTG	GTCTGAAGTA	TACAGCAGAG	GCTTGAAGGG	CATCTCGGAG	13580
	AGAAAAACAG	GCARATGAT	TAAAGAAAGT	GAAGAAAGAA	AAAGTGTAG	ATGGGAATTT	TCTTGTGAC	13650
	ATTTTATGCT	CCCAACCCAC	AGCTCAGATG	GTAGATGTGT	GTGAGAGCTA	ATGGACAGCA	CATATGACGA	13720
	AAAGCGAGC	CCATCTCTCT	AGAAAGCTGT	GTAACTGTGG	TATGTGGCAC	AGCTGATGGA	AAAGAGGTG	13790
45	TGTTGTGTAAT	TTTTTTTTCT	GAGAAACCTG	ACTGGAAGCA	ATAAGTTTGT	CTTCTTACAG	CATATACACG	13860
	AGCAGATTCT	AGGTGAAGGA	GGAGACAGAT	GCAAAACACA	CCACACACAG	AAATAAAACA	AAAGACTCAA	13930
	AGGGAAGGGA	GTGAACTGTT	CCCTGGTTTT	GTGTTGGGGA	AGGACACACA	GGGAGGCGGA	TGAACACAGT	14000
	GAGGCAACGG	GATTTGCTTT	CAGTGCAGAG	AACTCTGGTC	TGCTTGAGGC	CAGAGTAAAA	CCCACTATTCT	14070
50	TGTGGGCTGT	TGTGCAGCTG	ATTATTATTA	GAAGCTCTGC	AGAGCTCTGC	CAATCTCTGC	AGATCTCTGC	14140
	TCTCTCTAAC	CTCTGAGAG	GTAGAGGAGG	AAAGCTGCTA	GGGAGGAGAC	GGCCTTTGGT	CCACGAGCTG	14210
	GCAGAGGGCA	TGCATGATTT	CAGCTGTGCC	TCTGTCTCCG	GGGCTCTTGC	TCTGCCCAG	AGCCCCACAC	14280
	AAGTACAGCC	CATAGGCTCA	GGGTGAGCCG	GAGCCCAAGG	GTCTGTGGG	GATGGCTGTG	ATGAGAGAAA	14350
	TGACAGCTCT	TGCACACTTT	GGGAGAGTCC	TACCAGAGCT	TGCAAGAAAG	TGCATGTGAA	ACTGACAGCC	14420
55	AGAGCCATCT	CTCAAGAGAA	CCACGCTGAA	AGTATGGG	AGACCTGTCT	CAATCTCTCA	TGCTGGCTTC	14490
	TTTTTGGGCT	TGCTCAGAG	CCACATCTGA	GTGAGAGCA	CTCTGGAAG	CTTCTCTGGA	AGAGAGCTG	14560
	TTTGCATGGA	AGTCTCTACA	ATGTCTCTGT	TCTCCAGT	AAATTCACAT	CTGAAGTAGC	CAGACATATT	14630
	CACGGTCTTT	ATTTACCAT	TCCAGTGTTC	CAGGACGGGG	GACTTGCAC	AGCAAGTCA	GACACTGTCC	14700
	AAATACAGGG	CTAAGGAGAT	ATTATGCATC	ACAAAACCTG	CTCTGCCATT	AAACATTTAT	CAAGAATTTT	14770
60	TGAGAGATG	TTAATGGCA	CAAAAAGTTT	ATTTCAATGT	ACAGATGTTT	AGACATGGAT	GTAAGAGAAC	14840
	ACACGCCATC	AGGCTGGCCT	GAATGTCTAT	TGTGTTCTAT	TTTGGACATG	GCATCATATG	GGCAGTAGAT	14910
	GGTGTGAGG	CTCTCAGAG	CTCTGAGAG	AAAGCTGCTA	GGGAGGAGAC	CTCTGCTCT	TGGAGAGGCT	14980
	CTGCACTCA	CTGGAGCCTC	TGTAGCTGTG	TGCCACTGG	CTCTCTCATC	CTCTGAGATG	TAACACAGTG	15050
	AGATTTCCCA	GGCCCCAATC	AGTGTCTTCC	CACAAAACAC	GTAGTTCACA	CTGTGTCTCA	CTCGAGGAGC	15120
	GGCCGGGAG	CAGGGCTCCA	CAGTTTATTA	TGTGTTTTTG	GCTGAGTATG	GTGAGATGAT	CATCAGGGCA	15190
65	GATGATGATG	GCACAAACAC	GGCCGTCGGA	GGTTTGGATA	CATCTCAACT	CATCAGACAG	GTCTGTGGTG	15260
	AGTTTGGTCA	TGCAGAGCT	GGATGGGATG	TAGCATTTGG	AGTCTCATGA	GTGAGCACCC	AGCCCTCTTG	15330
	GGCTGCAGG	GCAGGAGAG	GCAGGACAG	GTGAGGAG	GAGGCGAG	GAAGGACAG	GGCTCTTGG	15400
	CACGAGAGGG	CTGTGGTGTG	GGGACGGCAC	CTGTGTCTGA	CATCTCCCCC	TGTGCTCAAG	CTATGCCCCG	15470
	TTTGGGCTCT	GGGCGTGAAG	TGTCACAGCC	GGCGGCTTCA	AGGCTGGGAG	CAACCTGTGT	CGCAAACTG	15540
70	AGAGTTCAGA	GTTCAGGAGG	TGTGTGGCGA	AGTATGTGTG	TGTGTTGTG	GGGCTGTGTC	TGGTAGGCTG	15610
	ATGCTGAGTC	GTGACAGCTA	AGAGTGCAGT	TGTACGATA	TACAGCTGTG	CAGCTAGATG	GTGTTGATGT	15680
	TGACATGAGA	CATGAGAGCT	GTGTGTCAGA	TGTGAGAGT	GGGACATG	CTGTGACATG	CTGTGACATG	15750
	CTGCACTGAT	ATGTTGGTCT	GTGACAGACT	GTGTGGGCAT	TCACGTGAGG	TGCTGAGTGT	TGGGTGTGCA	15820
	GTGTGAGTAG	CATGTGTGCA	CATAACATGT	ATTGAGGGGT	CTCTGTGTTT	ACCCGCTGAT	GTCTCTCAGA	15890
75	CCAGTGGCAC	TCCTTTACAGG	ATGAGACGGG	GTCCAGGGCT	TTGCTGGGCT	GAGGCTGCTG	TGCTGAGGCT	15960
	CTGAGGGCAT	TGTCCCATCT	GGGCACTGCG	GTCCACTCCG	TCTGCTGTGT	GGCTCTGTCT	CCCATCTCTC	16030
	TCTCCGTGG	GCATTTCAT	CTCCTCCACT	CCCTCTCTCC	TGTGGGAGTC	GGCTGTGCTAT	CCGCTCTCTT	16100

GTGGGACATCT	GGGTCCACCT	CGCCCTCTGTG	TGGGACATTG	GTGCACTCC	CTCTGCTGGT	TGCTTCTGTT	16246
CTTGCGCGAG	CGAGGATCT	CGAGGATCT	CGAGGATCT	TGACTTCCGC	AGGGTGTGTT	GGAGTGTGTT	16247
GTGGGCGGAG	AGGCGGATTT	TCACGTGGAA	GAGGAGATAGT	TTCTTGTCAA	AAATGTTCCCT	TTTCTTGTTT	16380
CATCTGAATG	GATGATTAAG	CAAAAGATAA	AAACTTAAAA	TCCGACAGAG	GTTTCTACCG	TTTCTACGCT	16450
TTTCTTGCGC	CGATTCAGTG	AACAGCCTCC	AGACGGTGTG	CACCAACATC	TACAGAGATC	TCTGCTGCA	16520
GGCGTACGGG	TGAGCGGCCA	CCAAGGGGTG	CAGGCGCAGC	CTCAGGGAC	CTCGAGCTCT	CGGCTACCT	16590
CTGACCCGGG	GGTTCACTCT	GGACATCTGT	GGTTTATGCG	GCAGGAAGAT	TCTTACGTTT	TGAGTGTGTC	16660
TCTTACCTGT	CGACACCTGT	GTCTGCTGTG	GTCTGCTGTG	AGACATCTGT	CTCATCTGTT	GGGTAGTGTG	16730
AGAGCGCGGT	GTGCGCCGAG	GTGCTCCGAG	TGTGCTCTGT	CACGACGCT	GGAGACTCAT	GGAGCGATCA	16800
CCAGGCGCAG	AGGGCGCATG	GGTAAAGAGA	GGTTTATGCG	GAGCTCTTACG	AGAGGAGGCT	GGGAAGATGT	16870
CTGAACAGTA	GATGGGAGAT	CAGATGCCCG	GAGGATTGCG	GAGGCTTACG	AGAGGGCGCG	AGGTGGGTGC	16940
AGGTGAGGGT	CGTGGGCCCG	ACCCCGGGGA	AGGTGACGCA	GAGCTGTGCG	TCCCCACACA	CGCCGCGCAG	17010
CACCTGTGCT	CTGGGCTATG	GTGTGCTCCT	GGAACTGTTC	GTCTGCTGCG	TGCTCAGAGG	GTGTCCTGCT	17080
CACAAATCGA	CACCTTTATC	ACAGAGGAAA	GGGCAATCT	GTGGAGGCCA	CAGGCGCAGC	TCTTCCCTGC	17150
AGTGAAGGCA	TGGGTGTGCA	CAGCTCTCGC	GGTGTACCA	AGGGCGATCT	GGGACACACA	GGCTCGGGCC	17220
TCAGACTCTAA	CAGCCTCTCC	GAGCCACTGG	GAGCTGAATG	CAAGGAGGCG	GAGGCCCTCG	CCCATCAGG	17290
GTGAGGAAGG	AGTGTGAGCA	TTTGTGTATC	CCAGGGCGGA	GGCTGGCGGA	ATTACCTGCG	ACATCTGATG	17360
TGAATGAGG	TGCTGCTCTA	TGCTGGAAGC	CCAGCAAGGG	CTCAGGGAG	AGTTTTCCAT	TACAAGGCTC	17430
TACCATGAAG	ATGTTGTTTT	ACCCGAGTGC	TTGGCGCTTC	ATGCTCTGCG	AGGGAGGGCA	GAGGCCACAG	17500
TGCTGTITAC	CGCTTTTGCA	CCAGCTCCAG	AGGCTTGGGA	CGAGCTGTCT	TCAGTTCCAG	GGTGGCTGCG	17570
CTGACAGCG	CGCTCTCTCT	TGCTTCTCTC	CTCTGCTCTA	AACTTCTCCT	CGTTTGCACT	TCCTCTAGCC	17640
GTGGCTGGGC	CTCTGTGCA	GTGCTGTGAC	TCTTCTCCGG	AAACCTTGG	GGTGTGCTGG	ATACAGGCTC	17710
CACCTGAGGAC	TGGAGGTGTC	TGACACTGTG	GTTCACCTCA	GGGTCAGCT	GGGCTGCTTT	GGGCTCTCTT	17780
GGGCTCATCAT	GAGGTCCAGG	GAGTGTTTCC	AGGTGAAGAC	TCTTGGGAAA	CTCCAGGCGC	CATGTGACCT	17850
CGGCACTGCT	CTCTCCATAT	CTGTGTCAGT	CTTGTCTCTA	TTTCCCGACC	AGGGTCTCTA	GGTCCGAGGA	17920
GGTCCATGAT	AGGCTGCGGG	CGAGGCGCAG	GGATGTGCTT	TTTCCCGACC	CATGTGGGGA	CCCTTGGGTA	17990
GTGCTGTGAT	TGGGTGAGCT	TGAGGAGGCT	GGGCTCAGCA	GGGCTCAGCA	ACGACAGGTA	ACGACAGGTA	18060
AGGCACTGTG	AGGCGCCAGC	AATCCCTCTC	CTCTGAGGCA	GGAGTGGGAG	AACGAGGAGC	TGGCGCCGTA	18130
TTTCAAGGCCA	CGAGGCGTGC	AGTGGGCGAG	CGTGTGGTGG	TCCACGTGCG	CTGTGGGGCG	GGGCTGATT	18200
CTGACAGGCG	GGGCTCTGCG	CTTCTCTGCC	CGTGTGGCG	CGGCTCTGCG	ACGGGCTTGG	GGGTGACCTC	18270
CGGACCTCTA	GCAGGTGTGCT	ATTCTCTCCT	TTTGGAGAGC	GGGCTCTGCG	CGTGTGAGT	GTTCCTCTCT	18340
TGGCTCATGA	TGCTGTGCTG	GTGTGCTGCT	GTGTGCTGCT	GGCTTCTTAA	TGCTTCTTAA	AACTCTGCTA	18410
CGCTGCTTTC	CTCTGTGCTT	AAATGGGAAA	AGACATCTCC	ACCTCAGGAG	AGTTCTAGAG	AGGCTGAAC	18480
CGGCTGCTGT	CTGTTCACTG	TGTGATCTCA	GGTCACTTCA	GAAGTGGCTG	AGGAGTCTAG	TGAGACAGG	18550
TACATGGGGG	CTGACGGCAG	TGGGTGAGAT	GAGGTACACG	GGGGGCTCAG	GAGGTGGGTG	AGGCTGAGTA	18620
CGTGGGGGGG	TCAGGCTCAT	GGTGTAGATG	GGTGTAGATG	GGGCTCAGCG	AGAGGGTCTAG	ACACGAGTCA	18690
CGGAGGCTCT	GATCAACAGC	ACATATGAGC	ACATGTGAC	ATGTTGCTGT	CTAGTGTAGC	AGGCTCTGTG	18760
CACACCTGCC	CGACAGGCTG	CACAGGCTG	CGTGTGCTG	ATGGGCTGCT	ACAGGCTGCT	ACAGGCTGCT	18830
CACACCTGTA	CTGTCACAGC	TTTGGAGGCG	CGAGGCGAGA	GAGTCCCTTG	AGGCCAGGAG	TTTAAGAGCA	18900
GGCTGAGGAA	CATATGAGAA	CCCATCTCT	ATGAAAATA	AAATCAAAAT	AGACTGTGAC	ATGGTGTGCT	18970
GGGCTGTATG	TTTCAATCAT	TGGGAGGCTG	AAATGGGAGG	ATCATCTGAG	CCAGGAGGCT	GGAGCTGCA	19040
GTGAGCTGAG	TTTCAACACG	TGTACTGCG	CGTGGGTGAG	AGAGTGTAGAG	CCATCTCA	CACACAGAAA	19110
GAGAGTACAG	AAATGCGACT	TTTGGGAAG	AAACATTTAG	TAGAACTTGA	AGCTTACAC	AGAGCCAGAG	19180
TGGGTGTCTG	GTGTGCTGTG	AGATGAGTGT	ATGGGTCTCT	AGACATCTAC	CCGACAGCCA	CCGACAGCCA	19250
ACCACAGGAG	CGGCTGGCTC	AGAAGGGATG	CGCAGGAGCT	TGATATACGA	TGATATACGA	TTGTGCTGAC	19320
GAGGCGGAGG	ATTTCATGATA	AGTACTGTCT	GGTACACAG	GAACATGGA	TAAATGGAA	AGCTTAGAGG	19390
CTTCCCGGCA	ACAGGCGGCT	ATCAGAAACC	AGCATGGGCG	AGCATGGGCG	AGGATGAGC	TGCTTACGCC	19460
TCCACATGCG	TGTTCAATACA	GATGTGTGAC	AGAAACCGAG	TGTACTGTG	CACACACAGA	CACGACGCTA	19530
CTGCAACACA	CACAGGCTG	CACAGGCTG	CGTGTGCTG	GTGCTGTG	TGACATCTG	CCGATCTGCA	19600
AACTCATGAG	TGTGCATCTA	TGCACACACA	CAGGACCGCG	TGGGCTGCTG	CCACACACA	CAGCACGCTG	19670
CTGATTAGGA	GGCTTTTCTT	CTGACGCTGT	CGGCATCTCT	CTCAGGTTCT	ACGCTATGTT	GGTGCAGCTG	19740
CCATTTCATC	AGCAAGTTTG	GAGAAACCCC	ACATTTTTCT	TGGCGCTTAT	CTCTGACACG	GCTTCTCTCT	19810
GTACTCTCAT	CTTGAAACCG	GGAGACGAG	GTATGTGCG	GTGCTGGGCC	TACGTCGAGC	CAGTGGCTGC	19880
CTCTGTGTGT	TMTGTGTGTA	GGAGACGAG	TGAATCTGCG	CTTAGAAGAT	CTTCTACCTT	TGTGTGATCA	19950
GGAGTGGTT	TTAGCGGACT	CTGTGTGCTG	CGGCTCTGCT	CGGCTCTGCT	GGGCTCTGCT	GGGCTCTGCT	20020
TGTAAGGAGC	AGGAGTCTGT	TGGAAGTCTG	CATCTCTTCC	ACCTTGTCTCT	CGCTGGGGGA	GGGCTGGGGG	20090
GTGTGGTCTC	CTCTGTTTGC	CCCATGGTGG	GATTTGGGGG	CTTGGCTCTC	CTCTGTTTGC	CTCTGTTTGC	20160
GATTTGGGCTG	TCTCCGCTAC	ATGSCACTTA	GGGCGCTTGT	GCAAAACGAG	GCCAAGGGCT	TAGAGAGAGG	20230
GCAGGCCGAG	GCATCCGCTAC	CCCTCTCAGG	AGCAGAGGCC	GGGATCTACC	AGCAGACAGC	CGGCGCGCTG	20300
CCCTTGCTTC	CCAGTGCAGT	GGTCTGGGCC	TGGGACATCT	TGTCGAGCAT	CAGGAGAGAT	TGTGATCTGT	20370
CTGATCTTCA	AGCATCTGCT	AGCAGCTGCT	CGTCACTTCA	CAACATCTCA	CAACATCTCA	CTTCTCTGTT	20440
GTGGAATTTT	CACCTGGAAG	AGCCGAAGAA	AACTTTCTCT	TGTGATCTCT	TGTGATCTCT	GGGCTGGGAG	20510
AGCAGAGATG	GGAGCGACCC	CGCAGACGCT	CGGGTGTGGG	CAGCTTTCGG	GTGTCTCTGT	GGAGGGGAGG	20580
TGGGCTGGGC	CTGTGATCTC	TGAGTGGTGT	TTTTGCCCGA	GGATGTGCG	TGTGGGCGCA	GGGCGGCCGC	20650
GGGCTGCTGC	CTCTCGAGGC	GGTGAAGTGG	CTGTGGCGAC	AGCATTTCTG	GCTCAAGCTG	ACTGCAAGCC	20720
GTCTCACTCT	CGTGGCGCTG	CTGTGGCTGAG	TGAGGACAGG	TAGGCTGGG	TAGGCTGGG	TAGGCTGGG	20790
AGTGGCGAG	GGTGAAGTGT	TGAGTGGTGT	GTGTGGCTG	GGGCTGTGCG	CTCTTCTGGA	ACCGAGATCT	20860
GGGCTGGGCG	GTGACCTGTG	GGGCTGGGAG	CCACCTTGG	AGCCTTATGT	GATTAACGCT	TGGTGTCCCG	20930
AGGCCAGCGA	GGCTGGGAGG	GGGCTGGGAG	TCTTGAACCC	CTGCTTCCCA	CTTCAAGGCG	GGGCTGTCCC	21000
CATGCTTGGG	AGGCTTGTGA	CCGCTTGTCT	GTGCTGCTCT	ACAGGCTCTT	CTCTGGGCTG	TGCGCTGAGG	21070
CTTGGGCTGT	CTGAGGAGAT	TCTGTGCTG	CGGCGGCTG	CCAGCTGGCT	TGGGCTGGCT	CTTGTGTGCT	21140
CGTGGGCTGT	GGGCTGTGCT	TGAGTGGTGT	AGGTTGGGAG	CAGTGGGAG	CTCTTCTGGA	ACCGAGATCT	21210
TGAGGCGCA	CGGCTGGGAG	GGGCTGGGAG	GTGCTGGGAG	GGGCTGGGAG	GGGCTGGGAG	GGGCTGGGAG	21280
ATTTGGGCGC	CGGAGCTGAG	AGCCGAATGA	GTGCGPAGCT	CTGCGGAGAG	AGGCTGACTG	CTCTGGAGAG	21350
CTTAGGCAAC	CGGAGCTGAG	CGGCTGGGAG	CGGCTGGGAG	GGGCTGGGAG	GGGCTGGGAG	CGGCTGGGAG	21420
CTTAGGCAAC	GACACAGGCA	CGGCTGGGAG	CGGCTGGGAG	CGGCTGGGAG	CGGCTGGGAG	CGGCTGGGAG	21490
CTTAGGCGCG	ACCGCTGGGA	CGGCTGGGAG	CGGCTGGGAG	CGGCTGGGAG	CGGCTGGGAG	CGGCTGGGAG	21560
CGTGTGATGT	CGGCTGGGAG	CGGCTGGGAG	CGGCTGGGAG	CGGCTGGGAG	CGGCTGGGAG	CGGCTGGGAG	21630

	CACCTCCCCA	CAGGCTGGGG	CTGGGCTCCA	CCCCAGGGGC	AGCTTTTCTT	CACCAAGGAG	CCGGGCTTCCA	21700
	CTCCCCACAT	AGGAATAGTC	CATCCGCCGA	TTCCGCCATT	TTCCGCCCTT	CCCTTTGCCCT	CTTTTGGCTT	21770
	CCACCCCCAG	CATCCAGGTG	GAGACCTCTA	GAGGACCTCT	GGGAGCTCTG	GGGATTGTGA	GTGACCAAGG	21840
	GTGTGCCCTG	TACACAGSGC	AGGACCCCTG	ACCTTGATGG	GGGTCCCTGT	GGGTCAAAAT	GGGGGAGAGT	21910
5	GTGTGGGAG	TAAATTAFTG	AATATATGAG	TTTTTCAGTT	TTGAAAAAAA	TCTCATGTTT	GAATCTTAAT	21980
	GTGCATCTGA	TAGACACACG	TGTATGCAAT	TACAGAAACC	TGTAGTGAAT	CGGGGTGGTG	GTCAGTGGGG	22050
	GGCCATGGCC	TGGCTGTGCA	TTTACGGGAG	TCTATGAGTG	AATGGGGTTG	TGGTCTAGTG	GGGGCCATGG	22120
	CTTGCGCTGGG	CTCGGGAGGT	TTCTGATGCT	GTGAGGCAAG	AGGGGAGAGA	GGGTAGGGGA	TAGACACTGG	22190
10	GAGCCGCCAG	CTCGGAGAGC	ATACCACTAA	GTCCAGACCT	GAGAGGACAG	AGGGATGCTG	GGGGCCAGCG	22260
	TTGGCGGGCG	GGGATGATGG	AGGGGCTGGC	CAGGGTGGCA	GGGATGATGG	GGGCCCCACG	TGGGGTGGCA	22330
	GGGGTGATGG	GGGGGCTTGG	CTCGGGTGGC	GGGGAGAGTG	GGGAGAGCTG	CTGGGGCCCT	CTGCTCCCTT	22400
	GGCTCCACCC	TGCGAGCCGTG	GATCCGGATG	TGCTTCCCTC	GTGCACATCC	TCTGGGCACT	CAGCTTTTAT	22470
	GAGATGGGGG	GGGAGGGGGA	TGACACCATC	CTGTATAAAA	TCCAGGATTC	CTCCTCTCTA	ACGCCCCACG	22540
	TCAGGTGAAA	AGTCACATTC	CGGCTCTGGC	CATTCTCTTA	AGAGTAGACC	AGGATTTCTA	TCTCTGAAGG	22610
15	CTGGTAGGGG	TGGGGCATGS	GAGGGTGGGG	ACACAGGAGG	CTTCAGGGTG	GGGCTGGTGA	TGCTCTCTCA	22680
	TCCTCTTAAT	ATCTCCCAAT	CTCACTCTCT	ATCTCTTAT	CATCTCCGAG	TCTCATCTGT	CTTCTCTCTA	22750
	TCTCCGAGTC	TACATCTGCA	TCTCTTACG	ATCTCCAGT	CTCATCTCTT	ATCTCTTAT	CTCTAGCTCT	22820
	CATCCAGACT	TCATCTCCAG	GGGGGCTGCC	AGGCTCGCAG	TGGAGCTGGA	CATACGCTCT	TCTCTAGACT	22890
20	GAGGAGACTG	GAAAGATATG	AGAGAACAGG	AGGGCGGCT	CAGAGGGACG	CAGTCTTGGG	GTGAGAGAAC	22960
	AGCCCCCTCT	CAGAAGTTGG	CTTGGGCCAC	ACGAAACCGA	GGGCCCTGGG	TGAGTGGCT	CAGAGCCTTC	23030
	CACAGCTCC	CTGTGGGGG	CTTATGTGAT	GGCGGGTCT	TACTGAGTGC	ACCTTGGAGC	GGGCTCTGGG	23100
	TTTGAATGCA	CGCGCGAGCT	GCTCTGGTCT	GGGGTGGGGG	CTTATGCCCA	CTGGATATGG	CGTCATTAAT	23170
	TGCTCTGGTG	TCAGAGGAATG	CTGAGGTGAC	CGAGGCTAAT	GTGTATGGTG	GGCCCAAGTG	CACAGAGACT	23240
25	GTGCTAAATG	CACCTCTGGTG	CTTGGAGCCC	CCGTATAGGA	GTGCTGAGGA	AGGAGGGGCT	CTTGGACGCC	23310
	GGCTCTGGGG	GGCTCTTGGC	CTGCAAACTG	GAAAGGAGCG	GCCTGGGGCG	CGCTGGGGCG	ACGACCTCAA	23380
	GTGAGAGGTT	GGACGAGACA	GGGCGGGGAC	TTCCAGGAGC	CAGAGGGGCG	TGCTCAGGCA	CACCTGGGTT	23450
	GAAGTACAG	ACCAAGGCT	CAGGCAATG	CTGACTTACG	AGGCTTTCAG	CACTCTCTCA	CACTCTCTCA	23520
	CTCCGGGTGT	TTTTTGTGTA	AATTTTACTC	AGGATTACTT	ATATTTTTTG	CTAAAGTATT	AGACCTTTAA	23590
30	AAAAGGTTAT	TGCTTTTGATA	TGGCTTAACT	CACATAAGCA	CTACTTTTAT	TGCTCTGTTT	TATTTTATAT	23660
	TATTATTATT	ATTAGAGAGTG	GTGTCTACTC	TGTCACCCAG	GTGTGTAATG	CAGTGGCACA	CTTCCGACTC	23730
	GTGTGAGCCG	CAAAACCCCA	GGCTCAAGTG	ATCTCTCCGG	CTCAGCTTCC	CAGAGTGTCT	GGATTACAGG	23800
	TGTGAGCCAC	TGGCCTTGGC	TGGCACTTTT	AAAACCAACT	ATGTAAAGTG	AGGTCACAGT	CTTTCCACAC	23870
35	CTGTCTCCCG	ATGATGTTGG	GAGGCCAGG	CAGAGGATCT	GTCTGAGGCG	AGGAGGTGTA	GAGCAGGTCT	23940
	GGTAACATG	GGAGACCCCA	CTCTTACAAA	AAATGCARAA	AGTTATCCGG	GGCTGGGGCT	CAGCATCTGT	24010
	AGTCCCAAGT	GCTCGGGAGG	CTGAGTGGGA	GGATCGCTTG	AGCCCGGAG	GTCAATGGCT	CAGTAGGCTG	24080
	TGATTGTACC	ATCGCACTTC	AGCCTGGGGA	ACAGAGTGAG	ACCCTGTCTC	AAAAAATAAA	AAAAAATAAA	24150
	AAGAGAGAAG	AAGAGAGAGG	AAGAGAGAGG	AAGAGAGAGG	AAGAGAGAGG	AAGAGAGAGG	AAGAGAGAGG	24220
	AAGAGAGGCT	GCTAGTGCTT	AGGTAGACTG	TCBAATCTCA	GAGCAAAAGT	AGTTTAAAGG	AGTTTAAAGG	24290
40	GGGAGAAAA	AGCCAGAGTC	TTTGGACTTC	CTTAGGCTCT	AGCATTCATCT	CAGACAGAGT	CTTCTCCACAG	24360
	ACAAGGCTGT	ATGAGGCGAG	TGAGTTCAAA	GCGAGAAGGG	AGGAGAAGCA	GAGCAGGGTG	GAGGCTGTGG	24430
	GTGACACAGG	CAAGGACCCC	TGAAGGGGAG	TGGTTGTTT	CTGCGCTCAG	CCCCACGCTC	CTGCCGGCTC	24500
	TGCACTCTGT	CTAACCGTGC	ATGTTGGTGC	CAGTGCCOCA	CTCGGAGAGG	ATGCTGTGCA	GGGGGCTTGC	24570
	CAACTCTTTG	TGGGTTTCAG	AAAGCCCCAG	CAGTCTGGCG	AGGCACAATT	ACAGCCCTCT	CCCAAGATGG	24640
45	CCGACGTCT	TCTCTGGGAA	CTGTGTGAAT	TGTCACCCCG	AAGGCCAGAG	CTGCTGAGAG	CTGAGCTGAG	24710
	AAATACGGGT	CGCATCTCAG	CGATCTTAAG	GTATCTCTGG	ATTATCTGGT	GGGCTTGATA	TGGCCACAAG	24780
	GGTCCCTAGA	AGTGAAGAGG	GGAGGACAGG	GAGAGTCAGA	GAGGGGACGT	GAGAAGGACG	ATGGGCCACT	24850
	GGGCTGTTTG	AGATGGAGGA	GGGGGCTCCC	AGCCAAAGAA	TGGGGGACAG	CGCTCTCATG	TGGAAAAAGC	24920
	AGCAATCTCT	CGCGGCTCTG	AGGGCACAAG	GCCTTGCCCA	CGCTCTGATT	TCAGGCGAGT	GGGACCTGTT	24990
50	TCAGTCTTCC	GGGCTCCAGA	GCTGTAAAGT	GATGGGTTTG	TGTTCCAGCA	GTTCAGCTGA	GTTCAGCTGA	25060
	ACAGACGAAA	ATGGAATAGC	AGTACAGGGA	ATGGAATACA	GGGACAGTTC	TCAGAGTGAG	TCACAGCCCA	25138
	CCCTGTGG							25138

Example 5

- 55 Comparison of the above-described genomic hTC sequence and the sequence of the hTC cDNA (Fig. 6; corresponding to SEQ ID NO 2) made it possible to elucidate the exon-intron structure of the hTC gene. The genomic organization of the hTC gene is illustrated diagrammatically in Fig. 7. The coding region of the hTC gene is composed of 16 exons which vary in size between 62 bp and 1354 bp (see Table 1)
- 60 Exon 1 contains the translation start codon ATG. The translation stop codon TGA and the 3'-untranslated region lie on exon 16 (Fig. 8). No possible polyadenylation signal (AATAAA) was found either in exon 16 or in the 3195 bp of the following

3'-flanking region. The exon-intron transitions were determined on the basis of the consensus sequence

	5'-Exon				Intron				3'-Exon			
5	Pre-mRNA	A/C	A	G		G	T	A/G	A	...	N C	A G G
	Frequency (%)	70	60	80		100	100	95	70		80	100 100 60

and listed in Table 1. With the exception of the 5' splice site between exon 15 and intron 15, all the exon-intron transitions are in accord with the published (Shapiro and Senapathy, 1987) splice consensus sequence. The sizes of the introns are between 104 bp and 8616 bp. Since only part of intron 6 was isolated, it is not possible to determine the precise length of the hTC gene. Based on the part sequence of ~4660 bp, which was obtained from intron 6, the minimum size of the hTERT gene is 37 kb.

0958246.002700

Introns 1-5 and the 5' region of intron 6, are contained in contig 1:

Intron 1: bp 11493-11596 (SEQ ID NO 4);

Intron 2: bp 12951-21566 (SEQ ID NO 5);

Intron 3: bp 21763-23851 (SEQ ID NO 6);

5 Intron 4: bp 24033-24719 (SEQ ID NO 7);

Intron 5: bp 24900-25393 (SEQ ID NO 8);

5' region of intron 6: bp 25550-26414 (SEQ ID NO 9).

The 3' region of intron 6, and introns 7-15, are located in contig 2 at the following positions:

10 3' region of intron 6: bp 1-3782 (SEQ ID NO 10);

Intron 7: bp 3879-4858 (SEQ ID NO 11);

Intron 8: bp 4945-7429 (SEQ ID NO 12);

Intron 9: bp 7544-9527 (SEQ ID NO 13);

i5 Intron 10: bp 9600-11470 (SEQ ID NO 14);

Intron 11: bp 11660-15460 (SEQ ID NO 15);

Intron 12: bp 15588-16467 (SEQ ID NO 16);

Intron 13: bp 16530-19715 (SEQ ID NO 17);

Intron 14: 19841-20621 (SEQ ID NO 18);

20 Intron 15: 20760-21295 (SEQ ID NO 19).

The 3'-untranscribed region is also located in contig 2 at position 21960-25138 (SEQ ID NO 20).

25 The individual sequences of the abovementioned introns are as follows:

001260.9428560

Intron 1 (SEQ ID NO 4)

GTGGGCTCCCCGGGTGCGGCGTCCGGCTGGGGTTGAGGGCGGCCGGGGGGAACAGCGACATCGGAGAGCAGCGCAGG
CGACTCAGGGGCGCTCCCCGGCAG

5 Intron 2 (SEQ ID NO 5)

GTGAGGAGTGTGGCCGTGAGGGGCCAGGCCCCAGAGCTGAATGCAGTAGGGGCTCAGAAAAGGGGGCAGGCAGAGCC
CTGGTCTCTCTGCTCCATCGTCAGTGGGCACACGTGGCTTTTCGCTCAGGACGTGAGTGACACGGTGATCTCTGCC
TCTGCTCTCCCTCTGCTCCGTTTGCATAACTTACGAGGTTACCTTCACGTTTGTATGGACACGCGGTTCCAGCGCC
CGAGGCCCAGAGCAGTGAACAGAGGAGGCTGGGCGCGCAGTGGAGCGGGTTGCCGCAATGGGGAGAAGTGTCTGGAAG
CACAGACGCTCTGGCGAGGGTGCCCTCAGAGTTACCTATAAATCTCTTCGCAATTTCAAGGGTGGGAATGAGAGGTGGGA
CGAGAACCCCTCTCTCTGGGGGTGGGAGGTAAGGGTTTTCAGGTGCACGTGGTCAGCCANTATGAGGTTTGTGTTTA
AGATTAAATTGTGTGTGACGGCCAGGTGCGTGGCTCAGCGCGGTAAATCCAGCAGCTTTGGGAAGCTGAGGCAGGTGGA
TCACCTGAGGTGAGGAGTTTGAGACAGGCTGACCAACATGGTAAACCTATCTGTAATAAAATCAAAAATAGCTG
GGCATGTGGTGTGTGCTCTGAATCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCACTTGAACCCAGGAGGCGGAGGC
TSCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCTGGGCGCAAGAGTGAACCTCTGCTTTAAAAAAGAAAGTGT
CGTGATTGTGCCAGGACAGGGTAGAGGGAGGAGATAAGACTGTTCTCCAGCAGAGATCTGGTCCCATCTTTAGGAT
GAAGAGGCGACATGGGAGCGAGAGCAGCAGATGGCTCCACCTCTGAGGAAGGACAGTGTGTTGGGTGTTCCAGGG
ATGGTGTCTGCTGGGCCCTGCCGTGTCGCCACCTGTTTTCTGGATTGATGTTGAGGAACCTCCGCTCCAGCCCCCTTT
TGCTGCCAGTCTGCCAGGCCCTACCGTGGCAGCTAGAAGAACTCCGATTTACCCCTCCGCCAAGACTCCCAAGAC
ATGTAAAGACTCCGGCCATGCGAGCAAGGAGGGTGAACCTTCTGGGGCTCTTTTTTCTTTTCTTTTCTTTGCTGGC
AAAAGTCATATAACATGAGATTGGCACTCCTAACACCGTTTCTGTGTACAGTCGCAAAATGCTAACTCGCGGTGTTTA
CAGCAGATTGCTGAAATGCTGGGTCTTGCCTGACGTGAAGTCCCTACCCATCGACAGCGAGCTGCTCACACACTCTGTGC
GGCTCAGGTGACACAGCGCCGATGAGATAAGCGTATGCAACCGAGTTTGCTTTTGTGCTCCAGGTTCTCTGCTGAG
GAGAGTTTGAGTTCTGATCAGGACTCTGCCTGTCATTGCTGTTCTGACTTCAGATGAGGTACAAATCTGCCCTGG
CTTATGCGAGGAGTAGAGCGTGGTCCCGGGTGTCCCTGTCACTGCAAGGTGAGTGAGGCGTTGCCCCAGAGTGTCCCT
GTCACTGTAGGGTGAAGTAGGCGGGCCCGGGTGTCCCTGTCCCGTGACAGCGATTGAGGTGTGGCCCGGGTGT
CCCTGTCACTGTAGGGTGAAGTAGGCGCCATCCCGGGTGTCCCTGTCACTGTAGGGTGAAGTAGGCGTGTCCCGG
GTGTCCTGTCCCGTGCAGGTTGAGTAGGCACTGTCCCGGGTGTCCCTGTCACTGTAGGGTGAAGTAGGCGCGGTCT
CCGGGTGTCCCTCTCAGGTGTAGGGTGAAGTAGGCGCGGCCCGGGTGTCCCTGTCACTGTAGGGTGAAGTAGGCG
GTCCCTGGGTGTCCCTCCGAGTTAGGGTGAAGTAGGCACTGTCCCGGGTGTCCCTGTCACTGTAGGGTGAAGTAGG
CGCGGCCCGGGTGTCCCTCAGGTGAGGGTGAAGTAGGCGGTGTCCCTGGGTGTCCCTGTCTGTGAGGTGAAGT
GAGGCTCTGTCCCGAGGTGTCTTGGCGTTTGTCTCACTTGAGCTTGTCTCTGAATGTTTGTCTTTCTATAGCCACAGCT
GCGCGGTGTGCCCATGCTCGGTGAGTGGTGAGGCGAGGTGCTGTCGCCAAGCCTATCTTTCTGATGCTCGGCTCT
TCTTGGTCACTCTCGTTCCATTTTGTACGGGACACGGGACTCAGGCTCTCGCTCCCGGTGCCAGGCACTGCAG
CCACAGCTTCAGGTCCGCTTGCCTCTGTGGGCTGGCTGTCTCACCAGCTGCCCGCCACATGCATGTCGCAATACTCC
TCTCCAGCTTGTCTAATGCGAGGCTGGACTGGGCTGCTGTCTGCTGCCAGTGTGCTGCTGGAGACATCCAGAA
AGGGTTCTCTGTGCCCTGAAGGAAGCAAGTCAACCCAGGCCCTCACTGTCTGTCTTCCCAAGCTGCCCTCTCTGC
TTGGCCCTCTGGTGGGTGGCAACGCTTGTCACTTATCTGGGCACTGCCCGCTCATTGCTTAGGCTGGGCTCTGCT
CCAGTGGCCCTCAGATGGAATTGACGTCAGGCGAGGTCAGGTTGGAAGTCTCTGTCTGCTGCTGAGGACAGCAGTG
GAGGGCGGTGTCTCCGCGAGGCTTCTGTCAGACTTCCCTCTGGGTCTTAGTTTGAATTTCACTGATTTACCTGTAGG
TTCTATCTCTCCTATGTAATGCTTTTCTTGGTCTATTCTTCACTGCTTTCTAGTTCTTAGTTTGTATGCTCTTTC
CCTCTAAGTGTGCTTACCTGACCCCTGTGTTTGTGAGTGAAGTAACTCAACATCAGGCACTTCAAGTGTCTTAA
ATACTTCAAAGTGTTAATACTTCTTTAAGTATCTTATTCTGTGATTTTTTCTTTGTGCAGCGTGTGTTTGCAGTGA
AATCATTTGATATCAGTCACTTTTAAAGTATCTTTAGCTTATCTGTGATTCTCTTAGCAGTGAGTATTTTGAACACT
GCTTATGTCGAAGATATGTAGAGTATCAAGTACGTAGAGTATTTAAGTATCATTTATTATTGATTTCTAACTCAAGT
TGTGTAGTGGTCTGTATAATACCAATTATTTGAAGTTTGGGAGGCTGTGCTTGTGTCTAGTGTGTGCATGTTTCCAG
AACTGTCCATTGTAATTTGACATCTCTCAATAGTGGGCAATGATTTCACTATATCCAGCTTATTAAGGTCCGTGCA

[illegible]

GAGGTATGGAGTCCGGATGATGCAGGTCCGGGGTGAAGTTGCCAGGCCCTGCTGTGAGCTGGATGTGCTGTATCCGGATG
GTGCAGTCCGGGGTGAAGTCCGACAGCCCTGCTGTGAGCTGGATGTGCTGTATCCGGATGGTGCAGGTCTGGGGTGAAGT
CACCAGGCCCTGCGGTGAGCTGGTTGTGCGGTGTCGGGTGCTGAGTCCGGGGTGAAGTCCGACAGGCCCTCGGTGAGC
TGGATGTGCGCGTGTCCCGGTGCCGATGGTGCAGGTCCAGGGTGAAGTCCGTAGGCCCTTGGTGGGTGGATGTGCGGT
5 GTCCGGATGGTGCAGGTCTGGGGTGAAGTCCGACAGGCCCTTGGTGAAGTGGATGTGCGGTGCTGTGATGGTGCAGGTCTG
GGTGAGGTGCCAGGCCCTTGGTGGGTGGATGTGTGCTGCGGTGTCAGGTGGTGCAGGTCCGGGTGAGGTCCGACAGGCCCT
GCTGTGAGCTGGATGTGCGGTGCTGGATGGTGCAGGTCCGGGGTGAAGTCCGACAGGCCCTCGGTGAGCTGGATGTGGG
GTGTCCGGATGGTGCAGGTCCGGGGTGAAGTCCGACAGGCCCTGCGGTGAGTGGATATGCGGTGTCGGATGGTGCAGGT
10 CCGGGTGAAGTCCAGAGGCCCTGCGGTGAGTGGATGTGCGGTGCTGATGGTGCAGGTCCGGGGTGAAGTCCGACAGGCCCT
CCCTGCTGTGAGCTGGATGTGCTGTATCCGGATGGTGCAGGTCCGGGGTGAAGTCCGACAGGCCCTGAGTGGTGGATGT
TGCTGTATCCGGATGGTGCAGGTCCGGGTGAGGTCCGACAGGCCCTGCGGTGAGTGGATATGCGGTGTCGGATGGTGCAG
GGTCCGGGGTGAAGTCCAGAGGCCCTGCGGTGAGTGGATGTGCGGTGTCGGATGGTGCAGGTCTGGGGTGAAGTCCGAC
AGGCCCTGCTGTGAGCTGGATGTGCTGTATCCGGATGGTGCAGGTCCGGGGTGAAGTCCGACAGGCCCTCGGTGAGCTGG
ATGTGCTGTATCCGGATGGTGCAGGTCTGGGTGAGTGGTCCGACAGGCCCTGCGGTGAGTGGATATGCGGTGTCGGATGG
15 TGCAGGTCCGGGGTGAAGTCCGACAGGCCCTGCGGTGGGTGATGTGTGTTGCTGAGTGGTGCAGGTCCGGGGTGAAGT
GCCAGGCCCTCGGGTGAAGTGGATGTGGTGTCTGGATGTCAGGTCCGGGGTGAAGTCCGACAGGCCCTCGGTGAGC
TGGATATCCGGTGTCCCGGTGCCAATGGTGCAGGTCCAGGTGAGGTGCCAGGCCCTTGGTGGGTGGATGTGAGTGGCT
TGCCGGATGGTGCAGGTCTGGGGTGAAGTCCGACAGGCCCTTGGTGAAGTGGATGTGCGGTGTCGGATGGTGCAGGTCCG
GGTGAGGTCCAGAGGCCCTCGGTGATCTGGATGGCATGTCTCTCTGTTAAG

Intron 3 (SEQ ID NO 4)

GTACTGTATCCGACGCGAGGCCCTGCTTCTCGAAGTCTGGAACACAGGCCCGGCTCAGCATGCGCCTGTCTCCACT
TGCCGTGTGCTTCCCTGCTGTGAGCTCTGGGTGGGAGCCAGGGGCCCTGACAGGCCCTGGTCCAGGTGGATGTGCTG
CAAGGCTCTGACTGCTCGGAGCTCACGTCTCTTACTTGTAAAAACAGGAGTTTGTGCCAAGTGGTCTCTAGGGTTTGT
25 AAGCAGAAGGGATTAAATTAGATGAAACACTACCATAGCTCCCTTGCCCTTCCCTGGGATGTGGGTCTGATCTGCTC
TCTCTTTTTTTTTCTTTTTGAGATGGAGTCTCACTCTGTTGCCAGGCTGGAGTGCAGTGGCATTAATCTTGGCTCACT
GCACCTCCACCTCCTGGGTTTAAAGGATTACACAGCCCTCAGCCTCCTAAGTAGTGGGATTACAGGCACCTGCCACCAC
GCCTGGCTTAATTTTGTACTTTTAGGAGAGACGGGGTTTACCATGTTGGCCAGGCTGGTCTCGAATCATGACCTCAGG
30 TGATCAACCACCCTTGGCTTCCCAAGTGTGGGTTTACAGGCTAAGCCAGGCTGGCCAGGCCCTCGATCTCTTTAATT
CATGCTGTCTGTATGAATCTCAATCTATTGGATTAGGTCTATGAGAGGATAAATCCCAACCACTTGGGACTCACTG
CAGGAGACCACTGTGACAGGAGACCTGGGATAGGAGAGTTCCACCATGAGCTAATCTTAGTGGGTGATTTGAATG
GCTGTGAGATTGTCTGCAATGTTCGGCTGATGAGAGTGTGAGATTGTGACAGATTCAAGCTGGATTGTGATCTGAGGT
GGACGGGAGCGCTGTCTGGAGATGCCAGCTGGCTGAGGCCAGGCCATGTTATGCTTCTCCGTGTCCCGCCAGGC
TGACTGTGGAGGGCTTTAGTCAGAAGTACAGGGCTTCCCAAGCTCCCTGCACACTCGAGTCCCTGGGGGGCCTTGTGAG
35 ACCCATGCCCAATCAGGATGTCTGACAGAGGAGCTGGCAGCAGACCTCGTCCAGAGGTAACACAGCCTTGGGCTGGG
GACCCCGACGTGGTGTGGGGCACTTCTCTGCTATGGGGGAGGGTACAGGCTTTCCTCGTGGGAAACAAGTTAATACAC
AATGCACCTTACTTAGACTTTACAGCTATTAAATGGTGTGGCAGCCACATGCTATTGACCAAGTATTGGAAAGAAAT
TTAATTGGGTGACCGGAAGGAGCAGACAGACGTGGTGGTCCCCAAGATGCTCTTGTCTACTACTGGGACTGTGTCTG
40 CTCTGGGGGGCCTTGGAGGCCCTTCTCTGCTGACAGGAGTACCGTGCCTTTTCTACTCTGCTGGGCCTGGGCCTGCGGT
AGGCCACAGCTCCGGGACACCCGCGGCCAGTGTCCACGGAGTGGCAGGCTGTCAGGACAGATGCCCCAGGTCCAGGT
GTGGCGCTCCAGCCCGCTGCCGCCAGTGGTGGTTTGGGGGAAAGGCCAAGGCCAGAGTGTGACAGGACTGTGGG
CTCATGAGAGCTGATCTGCTCTTGGTGAAGTGTGCTGAGCAGCTCTCCGCTCTCTCATCTGAAGGATGTGGCT
CTTCTAGCTGGGGTCTGGCTGGGCGGCCAGCCTTGGCTACCCAGTGGCTATACAGAGGACAGGCACTCTGTGTGG
AGGGGATGGGTTCAGCTGGCCCTGAGTGCAGCCTGGGACAGGCTCCCTGGTGTGATGGTGGGACAGTCAACCTGGGG
45 GTGACCGCGGAGCTGGGCGTCCCAAGGTGAGCTATAGGACAGGCTTCCCAAGTGGCTGTAAGTGGGGCTCTCAG
AGGCTGCTGGGTGGATGGGTGGCCGCGGAGTGGCTGAGCTGTGCTGGCTGGTGGCTGGTGGCTGGTGGCTGGTGGCT
TGAGTCCGTTGGGCTGTGGCTTCCGTGAGCTTCCGCTAGTCTGTGTGCTGGCTGAGCAAGCTCTGTGAGGCGCTCT
CTATTGACAG

GTCTGGGCACTGCGCCTGACGGTGTGGGACGGACTCCCAACAGTGGGTCTCCCCTGGGCAATCACTGGGCTCATGACGC
GACAGACTCTTGCGCCCTGTGGGCGCTGGGGGAGAAAGACTGTGATGGGGCATGATGAGCTGTGTGCTTCTGGCAATCT
TGAGCTGGGCACTGCAGCTCTGGCAGAGCTGCTGATCTACGACAGCTGTCTACGTTTGACTGCGCGCGCTCTCTCGATT
GCGCAGTGCCTTTGTTTCATGATTGTCTAAAGTCTTCTCTGCGAGTTTGAATCTGAGGCGCAAGGAAAGGTCTCCCGCT
CTTTTAGAGGGGCGGCGGCTATTTGAGCCGCTGTCTCTGCCAGCTGCCCTCTGAGCTGCTGGTCTGAGGCGCAAGGAAAGC
TGAGCCGCTCTCTAGGAGACGCGGCGCTGTTTGAGCAGCGCGCCGCTGAGCGGCGCTCTACAGTCTGGGCTCTGTCCAGCT

GGCCCTGTGGCCCTTTGCAGATGTGGTCTGTCCACGTGGCCCTGTGGCTCTTTGCAGATGCCTGTAGCACTTGCTCGGC
TCTAGGGGACAGTCTGTCCACCOCATGAGGCTCAGAGACCTCTGGGCGAATTTCTCTGGCTCCAGGGGTGGGGGTGGAG
GTGGCTGGGCTGTGGGACCAGACCTCTGTCCCGGCGAGTGGGCAGCACTCTGGATCACAATATGCCATCCGGGCCA
CGGTGGGTCTGTGGGTGTGAGCCAGCTGGACCCACAGGTGGCCAGAGGAGACGTTCTGTGTGCACACACTCTGTCTAA
5 GCCCATGTGTGTCTGCAGAGACTCGGCCCGGCAGCCACGATGGCCCTGCATTCAGCCAGCCCGCACTTCATCACA
AACACTGACCCCAAAAGGGACGGAGGCTCTTGCCACGCTGGTCTGTCTGACACCCAGCGCTCACTCCCATGTG
TCTCCGCTGTCTTTCCGAG

Intron 8 (SEQ ID NO 12)

- 10 GTGAGTCAGGTGGCCAGGTGCCATTCCTCGGGTGGCTGGCGGGCTTGCGAGGGCTCTGTCTCACCTCTCTCTCGCC
CTTCCCACTGNCCTTCTGCCCCGGGCCACAGAGTCTCTCTTTCTGGCCCCCGCCCCCTCGGCTCTCTGGGCTGCAGGC
TCCCAGGCCCGGAAACATGGCTCGGCTTGCGGCAGCCGGAGCGGAGCAGGTGCCACACGAGGCTGGAAATGGCAAGC
GGGCTGTGGAGTTGCTCTCTGCTGGAGGACGAGGGGGGGGGGTGTGTCTGGGTCAAGTGTGGCGCAGAGCCTTTGAGCCT
CGAGCTTTGTCACTCCAAGTTACTACTGACGTGGACACCCGGCTCTCACAGCTTGATCTCTCTCTCCGATACAAAA
15 GGATTTTATCCGATTCTCATCTCTGTCTCCCTGTCTGTGTGACCCCGGAGGGCGGGGCTCTCTCTCTGTGTAGATT
CCCATCTGGAAGTGGGGGTGACCGTGTAGTTTGTCTCTCGGGGGGCTGTGGTGGCCATGGGGCAGGCGGCTGG
GAGAGCTGCCGTCAACAGCCACTGGGTGAGCCACACTCAGGTGGTAGAGCCACAGTGGCTGGTGCACACTCAGCTCT
CTGGATTTTAAGTAAACACACACCTCCCGGAGGACATCTGCTGCGACCTGTGTGTCTGGGGAGAGGTAGACAC
10 GAGGAAATTCGTGCACACTCAAGGTATCAGCAAGGTATCCGAGTCAGGTGGAAGTGGAGGCTCTCTCTGGGATC
GTCTCAGCGGATAAAGGACTGTGCACAGCTTCGGAAGCTTTTATTAAATAATAACTATTAAATATTGCAATTATAAGT
AATCACTAATGGTATCAGCAATTATAATATTATTAAAGTATAATTAGAAATATTAAAGTATACACAGCTTCTGAAAAA
20 CACAATTTGCACATGGCAGCAGGTGAATTTGGCCGAGGGACACGTGTGCACATGTGTAAAGCGCCCCAGGCCCCAC
AGATTCGCTGACAAAGTCACTCCCCAGAGAAGCCACACGGGCTCTCTCTGTGTGTGAATTTATTAAAGTGGATC
AAGTCACGTACCGTCCACGTGTGGCAGGCTTTGGGGAATGTGAGGTGATGACTGCGCTCATGCCGTGACAGACAGGA
25 GGTGACTGTGTCTGCTCTGCTTCCTTAGGACACGGACAGGCGCGAAGCTCTAGTCCCCATCGTGGTCAAGTTGGGCTCTGA
ATAAAACGCTTCAAAACCTGTTGCCCAAAACTAAGAACAGAGAGGTTTCCCATCCATGTGCTCACAGGGCGTA
TGTGCTTGCCTGTGACTCGCTGGGCTGGCCGACTCTAGAGTTGGTGGTGTGCTCTGTGCAAAAAGTGCACTCTCT
GCCCATCACTGTGATCTGCACTGACAGCAAGGAAGCTCTTTCTTTCTTTTCTTTTGTGAGACGGAGCTCA
30 CTGTTGTCTGCTGGGCTTGAGTGCAGTGGCGCATCTCAACTCACTGCAACCTCCGCTCCCGGGTCCAGCATTTCTC
CTGCCTCAGCTCCCGAGCAGCTGAGATTACAGGCACCCACCCCTCGGCTTGCTAATTTTGTATTTTATGATAGAGAG
GGGTTTTGCCATGTTGGCCAGGCTGGTCTCGAATCTCTGACCTCAGGTGATCCACCCACCTCGGCTCCCCAAGTGTG
35 GATTACAGGTGTGAGCCATCAGCCCGGAGCGGAAGCTCTTTTAAAGTGACCACTATAGCGCTTCCGAAAAATAAC
AGGTCTGTGTTTTTGCAGTAGGCTGCAAGCGTCTCTAGCAACAGAGGTGGGCTCTGTGGGCTCTGGGATGGCTAGGAG
CTGGGCTGGACCATGCTCTCTGTGTCACCTTTAGGTTCCAGGGGCTATTCTGTCTCACTGTTTGTCTGAAACGCA
40 CCGTTCGATCCTTGTGTGAAGTTCTGTCTCTGTGTGTCATGCTGAACTAGGGGCAAGGTTGTATCTCGTTGGGCG
GCAGCGGCTACATGTAGGTCATGAGTCTTTACCGTGGACAAATTCCTTGAIAAAAAAAAAAGGAGTCCGTTAAGCAT
TCACTCCGGGTCAAGTGTCTGTGTTCTGTGAATAAATCTTAAGATTAAAGAAACCTTAATGAAAGAAACCTTTGATGATTC
5 AAGCAGAGATGTGGTCAACACTGTGGCTGGATCTGTTTCAGCCCGCCAGTGCATGTGAAGTGGGGAGGAGGATTG
TTGTGTCAGAGTCTCATCTGTATGTTTCTGAGGTGTTTGGCGGCTGAATGTTAGACGTGTGTTGTTGTGTATGAGGT
45 CTGTGTCTGTGTGTGGCTCGTTTGAAGTGTAGCATGTGCAGACATGCCCTGCCGTCTCTCATCTGTGTCTTCCGCG
CCAG

Intron 9 (SEQ ID NO 13)

- 10 TGAAGCTCTCTTTCGCCAGGGGGCTGGGTGGCGTTGATTTCTTTGATGCATTCATGTGTAATATTCCTGSGC
45 CTGTGAGACCATGACTGCTCTTCTTGAGGAACAGCAAGGTTTCAGGCCCTTCTTGGTATGAAGCCGCGCGGAGGGG

TTGCACAGCCTGAGGACTGCGGGCTCCACGCAAGCTCTGTCCAGCGGCATGTCCAGAGGCTCAGGAGGCGG
 GAGGGCCGCTGCCCTGCATGATGAGCATGTGAATTCAACACCGAGGAAGCACACCAAGCTTCTGTACAGTCAACAGGTTCC
 CGTTAGGGTCTTGGGAGATGGGCTGGTGCAGCCTGAGGCCCCACATCTCCAGCAGGCGCCCTGCAGAGGTTGGCTTGA
 CTGGGCGCTCTTTCAGCCGCTTGCACATCCCACTTGCATGGGCTACACCCAGGACGACACACCTAAATATCTGCGCC
 AACCTAATGTGGTTCAACTCAGCTGGCTTTATTGACAGCAGTTACTTTTTTTTTTAACTTTAAGTTCTAGGGTAC
 ATGTGCACGACGTGCAGGTTAGTTACATATGATACATGTGCCATGTTGGTGTGTGCACCCATTAACTCATCATTTACA
 TTAGGTATATCTCTAATGCTATCCCTCCCACTCCCCCATCCCATGACAGGCGCTGGTGTGTGATGTTCCCAACCTG
 TGTCCAAGTGTCTCATTTGTCAGTTGCCATGCCACTGTGAGTGAGAACATGTGGTGTGTTGGTTTTCTTCTTCCATGCAATAGTT
 GCTCAGAGTGATGGTTTCTCAGCTTCTGTCATGTCCCTACAAAGGACATGAACATCATCTTTTTTATGACTGCATAGTATT
 CCGTGGTGTATATGTGCCACATTTTCTTAATCCAGTCTATCATCGATGGACATTTGGGTTGGTTGCAAGTCTTGTCTACT
 GTGAATAGTGGCCCAATAAACATACGTGTGCATGTGCTTTATAGCAGCATGATTATAATCCTTTGGGTATATACCCAG
 TAATGGGATGGCTGGGTCAAAATGTTATTCTAGTTCTAGATCCTTGAGGAATCACCACATGTCTTCCCAATGGTGA
 CTAGTTTACACTCCACCAACAGTGTAAAAGTGTCTTGGTGTGTGAGAGGATGTGGACAGCAGTATTTTTTTATAGAAA
 TAGGCTACTGAACAGGACGACCTTAGTGAAGGATCGGTGAGGAGCCTGCAGGACACAGCCATTCTCTCGAAGAC
 TCCGGGTTTTCTGTGCATCTTTTGAACCTAGCTCCAATTATAGCATGTACAGTGGATCAAGGTTCTTCTTCATTAA
 GGTTCAGTTCTAGATTGAATAAGTTTATGTAAACAGAAACAAAATTTCTGTACACACAACTTGCTCTGGGATTGGA
 GGAAAGTGTCTCAGCTGGGCGCACACTGGTCAGCCCTCTGGGACAGGATACCTCTGGGCCATGGTCTATGGGCGCTGG
 GCTTGGGCTGAGGGTCACACAGTGCACCATGCCAGCTTCTGTGGATAGGATCTGGGCTCGGATCATGCTGTAGGACC
 ACAGCTGCCATGCTGTTAAAGGCGACACGTGGCTCAGAGGGGGCGAGGTTCCAGGCGCCAGCTTCTTACCGCTTTCAG
 TTATTTTTCCTAAGAGTCTGAGAAGTGGGCGCGCTGATGGCCTTCGTTGCTTCAGCTGGCAGAGAATTGCACAA
 GCTGAGTGAACACTGAGTACTTATAATGAATGAGGAATGCTGTAGCAGTAACTGTAGAGAGCTCGCTCTGTGGAAA
 GAAATTTAAGTTTTTCAATTAACCGCTTGGAGAATGTTACTTTATTTATGGCTGTGAAATTTGTTGACATTCACTGCC
 TCGTAGACAGATACTACGTAAAAAGTGTAAAGTTAACTTGTCTGTATTTCCTTATTTTAG

Intron 10 (SEQ ID NO 14)

GTGAGGCGCGTGCGTGTGTCTGTGGGACCTCCACAGCCTGTGGGCTTGCAGTTGAGGCCCCCGTGTCTGCCCTCG
 CACCGCAGCGTGTCTCTGCGAAGTCTCTCTCTGCGCGTGTGGATCGCAAGAGCAGAGGCGCTGCGCGCTGCAC
 CAGGCGCTGGGGCGCAGGGGACCTTGGGAGGAGTGGTACCGTGCAGGCGCTGCTCTGACAGGACGACCCAGCTT
 ACACAGTGGTGAAGTGCAGGCGGTGACCTGGCTCTGCTGCTCTTTGGAAAGTCAAGAGTGGCGGCTCTGGGCGCCAG
 TGAGACCCCGCAGGAGCTGTGCACAGGCGCTGCAGGCGCGAGGCGGACGCTCTCCCGAGGCTGCACCTGAGCTCGGGA
 GAGCAGGAGTCTGTAGTGAGCTGGCCACAGCGTTGCTGCGGTCACTTCTGCGTGGGTTGTTTGGGATCGTGGG
 AGAATTTGATTTGCTGAGTGTGCTGCTCTGAAACACCGAGATGGCTAGGAGTGGGTTTCAAGTTGATTTTGTGAAT
 CAAACTAAATCAGGCACAGGGGACTGGCTCAGCACAGGGGATTGTCCAATGTGGTCCCCCTCAAGGGCGCCCCACAG
 AGCCGCTGGGCTGTTTTAAAGTGCATTTGACGAGGACGAGAAACCTTGAAGCTGTAAAGGGAACCTCAGAAAAATG
 TGGCGCGTCAAGGCTGGTTTTCAGGTGCTTTGCTGGGCTGTGTTGTGAACCCATTGAGCCGCGCTCCAAATCCACCC
 TCCAGGTCCACCTTCCAGGGCGCGCTGGGCTGGGGTATGCTGCGGCTTCTTGTGCTGCAGCGCGGAGCAAGCAGGC
 TGTGACATTTAAATCCACTAAGATTCACTCGGGGAGGCCAGGTCCTCAAGCAACTGAGGCTCAGGAGCTGTAGGCT
 GCTGAGGGACAGACGACGAGCGGGAACGCTGTTCTGTGGCAAGTTCTGAGGGTGTGGCAGGAGGTGGCTCAGA
 GTGTATGTTGGGCTCCACCGGGGGCAGAACTCTGTCTGTGATGATCGGCAGCCATGTAAACAGAAAGGGTGGCCACAG
 GGAGCTGGGAATGCACAGGGAGGCTGCGCAGCTGGCGAGGTCACAGGGCCAGGCGCAAGGAAGGCGAGGGGAGCAGCC
 GGGGCGACAGCAGAGGCGCGAGGAAGGAAGGGATGCCAGGCGCAGACGAGGCTTCCGCGACAGGGGGCTCTCTG
 AGCTGGGTGAGCGAGGCTCATGACTCGCGAGGGAACCTCTTGACGTGAACCTGACGACTGTGTTGGCCATCTCAGAG
 CCCAGCAGGTCCCGGCGCTGAGCAGGAACCTCAGAACCTCCCTTTGTCTAAAGCAGCAGAGTGGCTTCAAGGATCT
 AGGAGAAAAACAGGCAAAAGTCGTTGAGAACTCTTAAAGAAAGGTGGGATGTTGGCAATTTCTTGTCCAGATTTAGTCT
 GCGCGGACACAGATGATGCTCTATAACGGATGTGGTGTGGCATGGGACACATGATGAGCAATCAGAGGCGAC
 TGGGCTGCACCTCCCATCTGAGTCTGGCTGTCGGGGCTCAGGCGCAGGTTCTTGCAATGCTCACTACCTCTCTCGCC

GGGAGACAGGGAAGCACCCGAGTCTGGAGCAGGGCTGGGTCCAGGCTCCTCAGAGCTCCTGCCAGGCCAGCACCCCT
GCTCCAAATCACCACCTCTCTCGGGGTTTCCAAAGCATTAAACAAGGGTGTAGGTTACCTCCTGGGTGACGGCCCCGCA
TCTCGGGGCTGACATTGCCCTCTGCCTTAG

5 Intron 11 (SEQ ID NO 15)

GTGAGCGCACCTGGCCGAAGTGGAGCTCTGCCCGGCTGGGCGAGGTGCTGTCGAGGGCCGTGCGTCCACCTCTGCT
TCCGTGTGGGCGAGCGACTGCCAATCCCAAAGGGTCAGAGGCCACAGGGTGCCCTCGTCCACTCTGGGGTGAGCAGA
AATGCATCTTTCTGTGGGAGTGGGGTGCTCAACAACGGGAGCAGCTTTCTGTGCTATTTTGGTAAAGGAAATGGTGCAC
CAGAGCTGGGTGCACCTGAGGTGTCTTCAGAAAGCAGTCTGGATCCGAACCCAGAGCCCGGGCCCTGCTGGGCGTGAGT
10 CTCTCAAACCCGACACAGAGGGCCCTGCTGGGCATGATCCCTCTGAACCCGAGACCTGGGGCCCTGCTGGGCGTGAGT
CTCTCCGAACCCAGAGACTTCAGGGCCCTTTTGGGCGTGAGTCTCTCGCTGTGAGCCCCACACTCCAAGGCTCATCCAC
AGTCTACAGGATGCCATGAGTTCATGATCACGTGTGACCCACTCAGGGGACAGGGCCATGGTGTGGGGGGGCTCTACAA
AATCTGGGCTCTGTCTTCCCGAGGCCGAGAGCTCAAGGCCCGCTCTCAGGCTCAGACACAATGAATTGAAGATGGA
CAGAGTGCAGAAATCTGTGCTGTTTCTTTATGAATAAAAGATATCAACATTCCAGGACGGGCAGGTGGCTCACAACCT
15 ATAAATCCGACACTTTGGGAGGCCGAGGTGGGTGATCACTTTGAGGCCAGGAGTTTGAGGCCAACCTTAACCAATAGTGTG
AAATTCACATTTCTACTTAAAAAATACAAAAATAGCCTGGCCTGGTGGCACACGCCGTGAGTCCCCGCTATGCGGGAGGC
TGAGGCAGGAGAATCATTTGAACCCAGGAGGCAGAGGTTGCAGTGAGCCGAGATCACACCACTGCATCCAGCTGGGCA
ACAGAGTGAGACTTCATCTTAAAAAAGTATCAGCATTCCAAAACCATAGTGAGCAGGTGTTTTTATTCT
TGTCCCTGCATAAATTTACTGTGTGCTGTGCTAGAGCCCGGAACCTGGGGTGCTTCTCTGAAAGGCACACCTTCATGG
20 GAAGAGAAATAAGTGGTGAATGTTGTTAAACAGAGGTTTAAACTGGGTCTCTGTTGTTCTGAGTTAACAGTCAGAGAT
TGGACTTTGCTCTTCCAGAAATGCTCCCTGGGTTTGTCTCATGGGGGAGCAGGAGTGGACACCCCTCGTGTAGGGG
GAGCAGCAGGTGCAGACGCCCTCATGATGGGGAGTGGCAGGTGCAGACACCTTCTGAGATGGTCCAGAGATGTCCTGT
TTGCAGCTCCCTCCCAACAGGATGCCGGTCTCCTGTGCTCCCCACAGTCCCTGCTTCCCTCTCACAGGCTTACCTGGTC
CTGGCCCTCAACTGGCTTTGTCTGCATGATTTCCACATTTCTCGGGCTCCAGCAGCTCTTTCGCCCTCTCCAGGCGACCT
25 GCAGTGTGGCCATACCACTCAGCTGTGAACGTCTCACTGCTTATTTTGTCCCCATGAAATGATTTTTTAGGACAGGC
ACCCCTGGTTCAGCCTCTGGGCACAGCATCAGTGAATGTTATTGAAGGACAAAGGACAGACAAACAATCAGGAAATGG
GTTCTCTCTAAACACATTCGAAAGCCACAGAGGCTAGTGCAGAGTGGTGGGCTCAGGTCATCAGATGTGGTCCAATG
CCAGAATATTTCTGTGCTCCCAAGGGCCACTTGGTCAGAGTGTGTGCTTCAGAGAGTGGCTCTAAAGACTCAGACCTAGGAG
GCAGTGGTTCGCCATTAATCAGGGTGAACTCACATCCTCTGTGTCTGAAGTATACAGCAGAGGCTTGAAGGGCATCTGGGA
30 GAAGAAAACAGGCAAAATGATTAAAGAAAAGTGAAGGAGGAAAGTGGTAAGATGGGAATTTTCTGTCCAGATTTTAGTCT
TCCCAAAACCAAGCTCAGATGGTGAAGTGTGTCAGAACTGATGGACAGAAACATAGAACAAAACCGGAAGCCCTATCTCT
CTGAAACAGTGTGTTAATGTGTGATGTGGCCAGAGTGTGATGAAAGAGAGTGTGTGTGATTTTCTGTGAGAAAACT
GACTCGAAGCAATAGTTGTCTTTACAGCATATACAGAGCAGATTCTAGGTAGAAGAGGAGACATGTCAAACAAC
ACCAGCAACAGAAATAAAACAAAAGACTCAAGGGGAAGGAGGTGAACGTTCCCTGGTGTGGTGTGGGGAAGGACACAC
35 AGGAGCGCGATGAACCCAGTGAGGCAACGGGCACTTCTTTCATCCAGAGAACTCAGGCTGGCTGAGCCACAGTGAAGA
ATGGCCATTCCTGGAGCGTTTGTGCAGCTGATTTATTAAGGCCCTGTGAGGTCCTGCACATTCATCCTCTCACTTT
GTTCTCTTAACCACTTGAGAGGTAGAGGAGGAAAGGCTCCAGGCGAGCAGCGCCCTTGGTCACCCAGCTGGCAAGAGGC
ATGCATGATTTGCAGCTGGCTCCTGCTCGGGGCCCTTGTCTGCGGAGACCCACACAAGTCAGACCCATAGGCTC
AGGTTGAGCCGAGGCCCAAGGTGCTGTTGGGGATGGCTGAAAGAAGAAATGGAGCTGTGATGCACACTTGGGAAGGTC
40 CTACACAGCCGCTCAGAGAAATGATGTGAACTCAGACGAGACCCATCCCTCAAGAAACGACAGTGAACCTGATGGC
GAGACCTGTCCCATCCCTCATGCTGGCTCCTTTCTGGGCTTGCCAAAGAGCCAGTCAGGTTGAGGCAAGCTGGAAG
ACTCTTTGGAAGAGCAGCTTGTGTGATGGAAGTCTCACAATGTCTGTGCTTCCAGTAATTCACCTCTGGAAGTGA
CCAGACATTAATCAGGGCTCTATTATACATTTTCAAGTGTTCAGGCGAGGGGACTTGCACAGCAAGTCAGCAACCTGCC
CAATACAGGGGCTAAGCAGATATTATGCATCAGAAAATCTGCTCTGCCATTAACATTTTTCAAGAAATTTTGAAGAAT
45 GTTAAATGACAAAACGTTTATTTTCAATGTAGCAGTGTCAAAGCTGGATGTAAGAAGAACACACCCAGGAGGAGTCCGGC
TGAATGTCATGTGTGTTCTTTSAGCATGGACATACATGGGCAGTGAGTGGTGTGAGGCCCTGGAGGACATCGGTGG

0956246.092100

5

Intron 12 (SEQ ID NO 16)

10

Intron 13 (SEQ ID NO 17)

25

30

35

40

43

CCGGGACCTTAGGCTTATTTATTTGTTTAAAAACATTTCTGGGCTGGCTTCGGTTGTTGCTAAATGGGAAAAAGACATCC
 CACCTCAGCAGAGTTTACTGAGAGGCTGAAACCGGGTGCTGGCTTGACTGGTGATCTCAGGTCATTCCAGAAGTGGCT
 CAGGAAGTCAGTGAGACAGGTACATGGGGGGCTCAGGCAGTGGTGAGATGAGGTACACGGGGGGCTCAGGCAGTGGT
 GAGGCCAGGTACATGGGGGGCTCAGGCAGTGGTGAGATGAGGTACACGGGGGGCTCAGGCAGAGGGTCAGACAGGTAC
 5 ACGGGGGCTCTGATCACACGCACATATGAGCACATGTGCACATGTGCTGTTTCATGTAGCCAGGTCTGTGCACACCTGC
 CCCAAGTCCCAAGAGGTGAGAGGCCAAGATGGAGGCTGACAGGGCTGGCGGGTGGCTCACACCTGTAGTCCCGACAG
 CTTTGGAGGCGGAGGCCAGAGGATCCCTTGAGCCAGGAGTTTAAGACCAGCCTGAGCAACATAGGTAGAACCCCATCTC
 TATGAAAAATAAAAAACAAAATAGCTGAACATGTTGGTGCGCCTGTAGTTCCAATACTTGGGAGGCTGAAGTGGGAG
 GATCACTTGAGCCCAGGAGGTTGAAGCTGCAGCTGAGCTGAGATGACACCACTGACTGCAGCCTGGGTGACAGAGTGAGA
 10 GCCCATCTCAACAACAACAAAGAGCTGACAAATGCAGTTTCTTGGAAAGAACATTTAGTAGGAACCTTAACCTACACA
 CAGAAGCCAAAGTCGGTGCTCCGGTGTCAGTGAGATGAGATGATGGGTCTTCACACCATCACCCACAGACCCAGGGTTTATG
 CACCACAGGGGGGGTGGCTCAGAGGGGATGCGCAGGACGTTGATATACGATGACATCAAGGTTGTCTGACGAAGGGCAG
 GATTTCATGATAAGTACCTGCTGGTACACAAGGAACAATGGATAAACTGGAACCTTAGAGGCCCTCCCGGAACAGGGGGCT
 AATCAGAAAGCCAGCATGGGGGGCTGGCATCCAGGATGGAGCTGCTTCAGCCTCCACATGCGTGTTCATACAGATGGTGCA
 15 CAGAAACGCAGTGTACCTGTGCACACACAGACCGCAGCTACTCGCACACACAAGCACACACAGACATGCATGCATGC
 ATCCGTGTGTGTGCACCTGTGCCATGAGGAAACCATGCATGTGCATTATGCAAGCACACAGGCACCGGTGGGCCCAT
 GCCCACACCCACGAGCAGCGTCTGATTAGGAGGCCCTTCTCTGACGCTGTCCGCCATCTCTCAG

Intron 14 (WEQ ID NO 18)

CTATGTGCGAGGTGCTGGCTCAGTGGCAGCAGTGCCTGCCTGCTGGTGTAGTGTGTAGGAGACTGAGTGAATCTGGG
 CTTAGGAAGTCTTACCCCTTTCCGCATCAGGAAGTGGTTAAACCAACCACTGTACAGGCTCGCTCGCCCGCCCTCTCGT
 20 GSGGTGAGCAGAGCAGCATGGGGGGCTGGCATCCAGGATGCTGCTGGGAGCTGCCATCTTCCCACTTGTCTGCCTGGGGAA
 GLGCTGGGGGGCTGGTCTCTCTGTTTGCCCATGGTGGGATTGGGGGGCTGGCCTCTCTGTTCCTCTGTGGTGG
 GATTGGGCTGTCICCCCTCCATGGCACTTAGGGCCCTTGTGCAAAACCCAGGCCAAGGGCTTAGGAGGAGGCCAGGCCAG
 25 GCTACCCCAACCCCTCTCAGGAGCAGAGGCCGCGTATCACCACGACAGAGCCCGCGCCTCTCTGCTTCCCACTCACCG
 TCCTCTGCCCTGGACATTTGTCCAGCATCAGGGAGGTTTCTGATCCGTCTGAAATTCAAAGCCATGCGAACCCTGCGGT
 CCTGAGCTTAAACGCTTCTACTTTCTGTTCTTCTGTGTTGTGGAAATTTACCTGGAGAACGCCGAAGAAACATTTCTG
 TCGTGACTCTCTGCGGTGCTTGGTGGGACAGCCAGAGATGGAGCCACCCGCAGACCCTGAGTGTGGGCGATTTCCG
 30 GTGTCTCTCTGGGAGGGAGCTGGGCTGGGCTGTGACTCCTCAGCCTCTGTTTTCCTCCAG

Intron 15 (SEQ ID NO 19)

GCAAGTGTGGGTGAGGCCAGTGGGGGCCCACTGCCAGGGGTATCCTTGAACGCCCTGTGTGGGGCGAGCAGCCTC
 AGATGCTGCTGAAGTGCAGACGCCCGGGCTGACCTGGGGGCTGGAGCCACGCTGGCAGCCCTATGTGATTAAACG
 35 CTGGTGTCCCCAGGCCAGGAGCTGGCAGGCTGCCCACTTCTTGAACCCCTGCTTCCCATCTCAGGGCGATGCTGCC
 CCACGCTTGGGAGCCTTCTGACCCCTGACCTGTGTCTCTCACAGCCTCTTCCCTGGCTGCTGCCCTCTGAGCTCTCTGGGT
 CTTGAGACGATTTCTCTCCCGCCCGCGCTCCAGCTCACTGGGTGCTGCTGTGCTGCGCCCGGTGGAGGGGTGCTG
 TCCCTTCACTGAGGTTCCCAACAGCCAGGSCCAGAGTGCAGGCCCTGCTGCGCCGCGCACACAGCTCTTAGGAGGG
 TTGGAGGATGCCACCTCTGGCCCTCTCTGGAACGGAGTCTGATTTGGCCCCGAG

3'-untranscribed region (SEQ ID NO 20)

ATCTCATGTTTGAATCCTAATGTGCATGCAATAGACACCACTGTATGCAATTACAGAGGCTGTGATGAACGGGGTGGT
 GGTCACTGGGGGCCATGGCCTGGCTGTGCATTACGGAAGTCTATGAGTGAATGCGGGTGTGTGTCATGCGGGGCCCATG
 40 GCTCGGTGGCCTGGGAGGTTCTGATGCTGTAGGCAGGAGGGGAGGAGGATAGGGGATAGACAGTGGGAGCCCCCA
 CCTTGAAGACATAACAGTAAGTCCAGGCCCGAAGGCAGCAGGAGTGTGGGGGCCAAGTGGGGCGGGGATGATG
 45 GAGGGCTGGCCAGGGTGGCAGGATGATGGGGGCCAGCTGGGTGGCAGGSGTGATGGGGGCTGTGCTGGGTGG

09502246_092100

[illegible]

Characterization of the exons showed, interestingly, that the functionally important hTC protein domains which are described in our Patent Application PCT/EP/98/03469 are arranged on separate exons. The telomerase-characteristic T motif is located on exon 3. The RT (reverse transcriptase) motifs 1-7, which are important for the catalytic function of the telomerase, are located on the following exons: RT motifs 1 and 2 on exon 4, RT motif 4 on exon 9, RT motif 5 on exon 10, and RT motifs 6 and 7 on exon 11. RT motif 3 is shared by exons 5 and 6 (see Fig. 8).

Elucidation of the exon-intron structure of the hTC gene also shows that the four deletions or insertion variants of the hTC cDNA which were described in our Patent Application PCT/EP/98/03469, as well as three additional hTC insertion variants which are described in the literature (Kilian et al., 1997), in all probability represent alternative splicing products. As shown in Fig. 8, the splicing variants can be divided into two groups: deletion variants and insertion variants.

The hTC variants in the deletion group lack specific sequence segments. The 36 bp in-frame deletion in variant DEL1 in all probability results from using an alternative 3' splice acceptor sequence in exon 6, resulting in a part of RT motif 3 being lost. In variant DEL2, the normal 5' splice donor and 3' splice acceptor sequences of introns 6, 7 and 8 are not used. Instead exon 6 is fused directly to exon 9, resulting in a displacement arising in the open reading frame and a stop codon appearing in exon 10. Variant Del3 is a combination of variants 1 and 2.

The insertion variant group is characterized by the insertion of intron sequences which lead to premature cessation of translation. Instead of the 5' splice donor sequence of intron 5, which is normally used, use is made, in variant INS1, of an alternative, 3'-located splice site, resulting in the insertion of the first 38 bp from intron 4 between exon 4 and exon 5. The insertion, in variant INS2, of a region of the intron 11 sequence likewise results from using an alternative 5' splice donor sequence in intron 11. Since this variant was only described inadequately in the

literature (Kilian et al., 1997), it is not possible to determine the precise alternative 5' splice donor sequence in this variant. The insertion of intron 14 sequences between exon 14 and exon 15 in variant INS3 comes from using an alternative 3' splice acceptor sequence, resulting in the 3' part of intron 14 not being spliced.

5

The hTC variant INS4 (variante 4), which is described in our Patent Application PCT/EP/98/03469, is characterized by exon 15, and the 5' part region of exon 16, being replaced by the first 600 bp of intron 14. This variant can be attributed to the use of an alternative internal 5' splice donor sequence in intron 14 and an alternative 3' splice acceptor sequence in exon 16, resulting in an altered C terminus.

10

The *in vivo* generation of hTC protein variants which are probably non-functional and which could interfere with the function of the complete hTC protein constitutes a possible mechanism, in addition to transcription regulation, for controlling hTC protein function. The function of the hTC splicing variants is not yet known. Although most of these variants presumably encode proteins without reverse transcriptase activity, they could nevertheless play a crucial role as transdominant-negative telomerase regulators by, for example, competing for interaction with important binding partners.

15

20

The search for possible transcription factor binding sites was carried out using the „find pattern“ algorithm from the Genetics Computer Group (Madison, USA) GCG Sequence Analysis program package. This resulted in the identification of a variety of potential binding sites for transcription factors in the nucleotide sequence of intron 2, which binding sites are listed in Tab. 2. In addition, an Sp1 binding site was found in intron 1 (pos. 43), and a c-Myc binding site was found in the 5'-untranslated region (cDNA position 29-34, cf. Fig. 6).

25

09582246-002100

Example 6

In order to ascertain the start point(s) of hTC transcription in HL 60 cells, the 5' end of the hTC mRNA was determined by means of primer extension analysis.

5

2 µg of polyA⁺ RNA from HL-60 cells were denaturated at 65°C for 10 min. 1 µl of RNasin (30-40 U/ml) and 0.3-1 pmol of radioactively labelled primer (5'GTAAAGTTGTAGCTTACACTGGTTCTC 3'; 2.5-8x10⁵ cpm) were added for primer annealing, and the whole was incubated, at 37°C for 30 min, in a total volume of 20 µl. After the addition of 10 µl of 5xreverse transcriptase buffer (from Gibco-BRL), 2 µl of 10 mM dNTPs, 2 µl RNasin (see above), 5 µl of 0.1 M DTT (from Gibco-BRL) 2 µl of ThermoScript RT (15 U/µl; from Gibco-BRL) and 9 µl of DEPC-treated water, primer extension took place, at 58°C for 1 h, in a total volume [lacuna]. The reaction was stopped by adding 4 µl of 0.5 M EDTA, pH 8.0, and the

10 RNA was degraded, at 37°C for 30 min, after having added 1 µl of RNaseA (10 mg/ml). 2.5 µg of sheared calf thymus DNA and 100 µl of TE were then added, and the mixture was extracted once with 150 µl of phenol/chloroform (1:1). The DNA was precipitated, at -70°C for 45 min, after adding 15 µl of 3 M Na acetate and 450 µl of ethanol, and then centrifuged at 14,000 rpm for 15 min. The precipitate was

15 washed once with 70% ethanol, dried in air and dissolved in 8 µl of sequencing stop solution. After 5 min of denaturation at 80°C, the samples were loaded onto a 6% polyacrylamide gel and fractionated electrophoretically (Ausubel et al., 1987) (Fig. 5).

20

25 In this connection, a main transcription start site was identified which is located 1767 bp 5' of the ATG start codon of the hTC cDNA sequence (nucleotide position 3346 in Fig. 4). In addition to this, the nucleotide sequence around this main transcription start (TTA₋₁TTGT) represents an initiator element (Inr), which, in 6 out of 7 nucleotides, matches the consensus motif (PyPyA₋₁Na/tPyPy) (Smale, 1997) of

30 an initiator element.

It was not possible to identify any unambiguous TATA box in the immediate vicinity of the experimentally identified main transcription start, which means that the hTC promoter has probably to be classified in the family of TATA-less promoters (Smale, 1997). However, a potential TATA box from nucleotide position 1306 to nucleotide position 1311 (Fig. 4) was found by means of bioinformatics analysis. The subsidiary transcription starts which were additionally observed around the main transcription start have also been described in the case of other TATA-less promoters (Geng and Johnson, 1993), for example in the strongly regulated promoters of some cell cycle genes (Wick *et al.*, 1995).

Example 7

In addition to the start point of the hTC transcript which was described in Example 6 and identified in HL60 cells, a further transcription start region was also identified in HL60 cells. With the aid of RT-PCR analyses, the region of the hTC gene transcription start in HL60 cells was localized to bp -60 to bp -105.

The cDNA for this was synthesized using a First Strand cDNA Synthesis kit (Clontech), in accordance with the manufacturer's instructions, and employing 0.4 µg of HL60 cell polyA RNA (Clontech) and the gene-specific primer GSP13 (5'-CCTCCAAAGAGGTGGCTTCTTCGGC-3', cDNA position 920-897). In a final volume of 50 µl, 10 pmol dNTP mix were added to 1 µl of cDNA, and a PCR reaction was carried out in 1xPCR reaction buffer F (PCR-Optimizer kit from InVitrogen) and using one unit of platinum Taq DNA polymerase (from Gibco/BRL). 10 pmol of each of the 5' and 3' primers defined below were added as primers. The PCR was carried out in 3 steps. A two-minute denaturation at 94°C was followed by 36 PCR cycles in which the DNA was first of all denatured at 94°C for 45 sec and, after that, the primers were annealed, and the DNA chain was extended at 68°C for 5 min. The cycles were concluded by a chain extension at 68°C for 10 min. In all, six different 5' PCR primers (primer HTRT5B: 5'-CGCAGCCACTACCGCGAGGTGC-3', cDNA position 105 to 126; primer C5S:

5'-CTGCGTCCTGCTGCGCACGTGGAAGC-3', 5'-flanking region -49 to -23; primer PRO-TEST1: 5'-CTGCGGCGCGAGTTTCAGGCAG-3', 5'-flanking region -74 to -52; primer PRO-TEST2: 5'-CCAGCCCCCTCCCCTTCCTTCC-3', 5'-flanking region -112 to -91; primer PRO-TEST4: 5'-CCAGCTCCGCCTCCTCCGCGC-3', 5'-flanking region -191 to -171; primer RP-3A: 5'-CTAGGCCGATTCGACCTCTCTCC-3', 5'-flanking region -427 to -405) were combined with the 3' PCR primer C5Rback (5'-GTCCCAGGGCACGCACACCAG-3', cDNA position 245 to 225). Genomic DNA was also employed for the PCR, as a control, in addition to the Oligo dT- and GSP13-primed cDNAs. As Fig. 9 shows, a PCR product was only obtained with the primer combinations HTRT5B-C5Rback, C5S-C5Rback and PRO-TEST1-C5Rback, indicating that the start point for hTC transcription lies in the region between bp-60 and bp-105.

15 **Example 8**

Several extremely GC-rich regions, so-called CpG Islands, are located in the isolated 5'-flanking region, of about 11.2 kb in size, of the hTC gene. One CpG Island, having a GC content of > 70%, extends from bp - 1214 into intron 2. Two further GC-rich regions having a GC content of > 60% extend from bp -3872 to bp -3113 and from bp -5363 to bp -3941, respectively. The positions of the CpG Islands are shown graphically in Fig. 11.

The search for possible transcription factor binding sites was carried out using the "Find Pattern" algorithm from the Genetics Computer Group (Madison, USA) GCG Sequence Analysis program package. This resulted in the identification of a variety of potential binding sites in the region up to -900 bp upstream of the translation start codon ATG: five Sp1 binding sites, one c-Myc binding site, and one CCAC box (Fig. 10). In addition, a CCAAT box and a second c-Myc binding site were found at positions -1788 and -3995, respectively, of the 5'-flanking region.

Example 9

In order to analyse the activity of the hTC promoter, PCR amplification was used to generate four hTC promoter sequence segments of differing length, which segments were cloned into the Promega vector pGL2 5' in front of the luciferase reporter gene. The 8.5 kb SacI fragment which was subcloned from phage clone P12 was selected as the DNA source for the PCR amplification. In a final volume of 50 µl, 10 pmol of dNTP mix were added to 35 ng of this DNA, and a PCR reaction was carried out in 1xPCR reaction buffer (PCR-Optimizer kit from Invitrogen) and using one unit of platinum Taq DNA polymerase (from Gibco/BRL). In each case 20 pmol of the 5' and 3' primers which are defined below were added as primers. The PCR was carried out in three steps. A two-minute denaturation at 94°C was followed by 30 PCR cycles in which the DNA was first of all denaturated at 94°C for 45 sec, after which the primers were annealed, and the DNA chain was extended, at 68°C for 5 min. The cycles were concluded by a chain extension at 68°C for 10 min. The selected 3' PCR primer was in each case the primer PK-3A (5'-GCAAGCTTGACGCAGCGCTGCCTGAAACTCG-3', position -43 to -65), which primer recognizes a sequence region 42 bp upstream of the ATG START codon. A promoter fragment of 4051 bp in size (NPK8) was amplified by combining the PK-3A primers with the 5' PCR primer PK-5B (5'-CCAGATCTCTGGAACACAGAGTGGCAGTTTCC-3', position -4093 to -4070). Combining the pair of primers PK-3A and PK-5C (5'-CCAGATCTGCATGAAGTGTGTGGGGATTGCAG-3', position -3120 to -3096) led to the amplification of a promoter fragment of 3078 bp in size (NPK15). Use of the primer combination PK-3A and PK-5D (5'-GGAGATCTGATCTTGGCTTACTGCAGCCTCTG-3', position -2110 to -2087) amplified a promoter fragment of 2068 bp in size (NPK22). Finally, using the primer combination PK-3A and PK-5E (5'-GGAGATCTGTCTGGATTCTGGGAAGTCCTCA-3', position -1125 to -1102) led to the amplification of a promoter fragment of 1083 bp in size (NPK27).

0956246, 092100

The PK-3A primer contains a HindIII recognition sequence. The different 5' primers contain a BglII recognition sequence.

5 The resulting PCR products were purified using the Qiagen QIA quick spin PCR purification kit, in accordance with the manufacturer's instructions, and then digested with the restriction enzymes BglII and HindIII. The pGL2 promoter vector was digested with the same restriction enzymes, and the SV40 promoter contained in this vector was released and removed. The PCR promoter fragments ligated into the
10 vector, which was then transformed into competent DH5 α bacteria (from Gibco/BRL). DNA for the promoter activity analyses, which are described below, was isolated from transformed bacterial clones using the Qiagen plasmid kit.

Example 10

15 The activity of the hTC promoter was analysed in transient transfections in eukaryotic cells.

All the work with eukaryotic cells was carried out at a sterile workstation. CHO-K1 and HEK 293 cells were obtained from the American Type Culture collection.

20 CHO-K1 cells were kept in DMEM Nut Mix F-12 cell culture medium (from Gibco-BRL, order number: 21331-020) containing 0.15% streptomycin/penicillin, 2 mM glutamine and 10% FCS (from Gibco-BRL).

25 HEK 293 cells were cultured in DMOD cell culture medium (from Gibco-BRL, order number: 41965-039) containing 0.15% streptomycin/penicillin, 2 mM glutamine and 10% FCS (from Gibco-BRL).

30 CHO-K1 and HEK 293 cells were cultured at 37°C in a water-saturated atmosphere while being gassed with 5% CO₂. When the cell lawn was confluent, the medium was sucked off, after which the cells were washed with PBS (100 mM KH₂PO₄ pH

7.2; 150 mM NaCl) and released by adding a trypsin-EDTA solution (from Gibco-BRL). The trypsin was inactivated by adding medium and the cell count was determined using a Neubauer counting chamber in order to plate out the cells at the desired density.

5

For the transfection, in each case 2×10^5 HEK 293 cells were plated out, per well, in a 24-well cell culture plate. The HEK 293 medium was removed after 3 hours. For the transfection, up to 2.5 μ g of plasmid DNA, 1 μ g of a CMV β -Gal plasmid construct (from Stratagene, order numner: 200388), 200 μ l of serum-free medium and 10 μ l of transfection reagent (DOTAP from Boehringer Mannheim) were incubated at room temperature for 15 minutes and then dropped uniformly onto the HEK 293 cells. 1.5 ml of medium were added after 3 hours. The medium was changed after 20 hours. After a further 24 hours, the cells were harvested for determining the luciferase activity and the β -Gal activity. For this, the cells were lysed, at room temperature for 15 minutes, in the cell culture lysis reagent (25 mM Tris [pH 7.8] containing H_3PO_4 ; 2 mM CDTA; 2 mM DTT; 10% glycerol; 1% Triton X-100). Twenty μ l of this cell lysate were mixed with 100 μ l of luciferase assay buffer (20 mM Tricin; 1.07 mM $(MgCO_3)_4$ $Mg(OH)_2 \cdot 5H_2O$; 2.67 mM $MgSO_4$; 0.1 mM EDTA; 33.3 mM DTT; 270 μ M coenzyme A; 470 μ M luciferin, 530 μ M ATP), and the light generated by the luciferase was measured.

10

15

20

In order to measure the β -galactosidase activity, equal quantities of cell lysate and β -galactosidase assay buffer (100 mM sodium phosphate buffer, pH 7.3; 1 mM $MgCl_2$; 50 mM β -mercaptoethanol; 0.665 mg of ONPG/ml) were incubated at 37°C for at least 30 minutes or until a slight yellow coloration appeared. The reaction was stopped by adding 100 μ l of 1 M Na_2CO_3 , and the absorption was determined at 420 nm.

25

In order to analyse the hTC promoter, four hTC promoter sequence segments of differing length were cloned 5' in front of the luciferase reporter gene (cf. Example 9).

30

The relative luciferase activities of two independent transfections in HEK 293 cells, using the constructs NPK8, NPK15, NPK22 and NPK27, are plotted in Fig. 11. Each experiment was carried out in duplicate. The standard deviation has also been given.

5 The construct NPK 27 exhibits a luciferase activity which is 40 times higher than the basal activity of the promoterless luciferase control construct (pGL2-basic) and from 2 to 3 times higher than that of the SV40 promoter control construct (pGL2PRO). Interestingly, a luciferase activity which was from 2 to 3 times lower than that

10 obtained with the NPK 27 construct was observed in cells which were transfected with longer hTC promoter constructs (NPK8, NPK15, NPK22). Similar results were also observed in CHO cells (data not shown).

00582246,092100

References

- 5 Allsopp, R. C., Vazire, H., Patterson, C., Goldstein, S., Younglai, E.V., Futcher, A.B., Greider, C.W. und Harley, C.B. (1992). Telomere length predicts replicative capacity of human fibroblasts. *Proc. Natl. Acad. Sci.* **89**, 10114-10118.
- 10 Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (1987). *Current protocols in molecular biology*. Greene Publishing Associates and Wiley-Intersciences, New York.
- 15 Blasco, M. A., Rizen, M., Greider, C. W. und Hanahan, D. (1996). Differential regulation of telomerase activity and telomerase RNA during multistage tumorigenesis. *Nature Genetics* **12**, 200-204.
- 20 Broccoli, D., Young, J. W. und deLange, T. (1995). Telomerase activity in normal and malignant hematopoietic cells. *Proc. Natl. Acad. Sci.* **92**, 9082-9086.
- 25 Counter, C. M., Avilion, A. A., Lefevre, C. E., Stewart, N. G. Greider, C.W. Harley, C. B. und Bacchetti S. (1992). Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.* **11**, 1921-1929.
- 30 Feng, J., Funk, W. D., Wang, S.-S., Weinrich, S. L., Avilion, A.A., Chiu, C.-P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J., Le, S., West, M.D., Harley, C.B., Andrews, W.H., Greider, C.W. und Villeponteau, B. (1995). The RNA component of human telomerase. *Science* **269**, 1236-1241.
- 35 Geng, Y., and Johnson, L.F. (1993). Lack of an initiator element is responsible for multiple transcriptional initiation sites of the TATA less mouse thymidine synthase promoter. *Mol. Cell. Biol.* **14**:4894.
- Goldstein, S. (1990). Replicative senescence: The human fibroblast comes of age. *Science* **249**, 1129-1133.
- Harley, C.B., Futcher, A.B., Greider, C.W., 1990. Telomeres shorten during ageing of human fibroblasts. *Nature* **345**, 458-460.

- Hastie, N. D., Dempster, M., Dunlop, M. G., Thompson, A. M., Green, D.K. und Allshire, R.C. (1990). Telomere reduction in human colorectal carcinoma and with ageing. *Nature* **346**, 866-868.
- 5 Hiyama, K., Hirai, Y., Kyoizumi, S., Akiyama, M., Hiyama, E., Piatyszek, M.A., Shay, J.W., Ishioka, S. und Yamakido, M. (1995). Activation of telomerase in human lymphocytes and hematopoietic progenitor cells. *J. Immunol.* **155**, 3711-3715.
- 10 Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C. B., West, M.D., Ho, P.L.C., Coviello, G.M., Wright, W.E., Weinrich, S.L. und Shay, J.W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-2015.
- Latchman, D.S. (1991). Eukaryotic transcription factors. Academic Press Limited, London.
- 15 Lingner, J., Hughes, T.R., Shevchenko, A., Mann, M., Lundblad, V. und Cech T.R. (1997). Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* **276**: 561-567.
- Lundblad, V. und Szostak, J. W. (1989). A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* **57**, 633-643.
- 20 McClintock, B. (1941). The stability of broken ends of chromosomes in *Zea mays*. *Genetics* **26**, 234-282.
- Meyne, J., Ratliff, R. L. und Moyzis, R. K. (1989). Conservation of the human telomere sequence (TTAGGG)_n among vertebrates. *Proc. Natl. Acad. Sci.* **86**, 7049-7053.
- 25 Olovnikov, A. M. (1973). A theory of marginotomy. *J. Theor. Biol.* **41**, 181-190.
- Sandell, L. L. und Zakian, V. A. (1993). Loss of a yeast telomere: Arrest, recovery and chromosome loss. *Cell* **75**, 729-739.
- 30 Shapiro, M.B., Senapathy, P., 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucl. Acids Res.* **15**, 7155-7174.
- 35 Smale, S.T. and Baltimore, D. (1989). The „initiator“ as a transcription control element. *Cell* **57**:103-113.

Smale, S.T. (1997). Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochimica et Biophysica Acta* 1351, 73-88.

- 5 **Shay, J. W. (1997).** Telomerae and Cancer. Ciba Foundation Meeting: Telomeres and Telomerase. London.

- 10 **Vaziri, H., Dragowska, W., Allsopp, R. C., Thomas, T. E., Harley, C.B. und Landsdorp, P.M. (1994).** Evidence for a mitotic clock in human hematopoietic stem cells: Loss of telomeric DNA with age. *Proc. Natl. Acad. Sci.* **91**, 9857-9860.

- 15 **Wick, M., Härönen, R., Mumberg, D., Bürger, C., Olsen, B.R., Budarf, M.L., Apte, S. S. and Müller, R. (1995).** Structure of the human TIMP-3 gene and its cell-cycle-regulated promoter. *Biochemical Journal* 311, 549-554.

Zakian, V. A. (1995). Telomeres: Beginning to understand the end. *Science* **270**, 1601-1607.

00000000.00000000

SEQUENCE LISTING

	<110> Bayer AG	
5	<120> Regulatory DNA sequences from the 5i region of the gene for the human catalytic telomerase subunit, and their diagnostic and therapeutic use	
10	<130> LeA32805-Foreign Countries	
	<140>	
	<141>	
15	<160> 20	
	<170> PatentIn Vers. 2.0	
	<210> 1	
20	<211> 5126	
	<212> DNA	
	<213> Homo sapiens	
	<400> 1	
25	gagctcttgaa ccgtggaaac gaacatgacc ctgacctgcc tgettccctg ggtgggtgcaa 60	
	gggtaagttaa gtgtgtgtgca ggaatggccc atgtaaatata cagcactctg ctgagtgaggaa 120	
	ccgttctcttc catcattatt catcttcaacc cccaaggacct gaatgatctc agcaactctct 180	
	tcgggtgtgta caagccatga caaactcag tacaacacac agcttttca tagggccaca 240	
	gagcacaggc cactctcccg atattattaa agtcaggag agtcaggacct gctttccagc 300	
30	accagctcgg ggtgacacaa cgcgctgaac agtctgttcc tctagactag tagacctcgg 360	
	caggactcgc cccaatttct agggcctggg tctgtcttcc cgaggcgccc atctgcctcg 420	
	gagactcagc ctgggggtgcc acactgaggc cagccctgtc tccacacctc cgcctccagc 480	
	gctcagcttt cctccagacc ttccctaaacc ctgggtgggc cgtgttccag cgtcactatgc 540	
	tcactctgtc cactgtgtct gctctcagc agtgaactcg cactgttctc cctccatctg 600	
35	ggtgtctgtc tcttcccca acactcacat cggttagagg gaggagattc tgcgctctcc 660	
	agactggctc ctctgagcct gaacctggct cgtggccccc gatcagattc cttggctctc 720	
	ggctgcacgc tgacctccat ttccaggcgc tcccgtcttc ctgtcatctg ccggggcctg 780	
	ccgggtgtgt ctctgttttc tgtgtctctt tccaagttca gctcggttgc tctctgcccc 840	
	ctagggtctc ggggttttta taggcatagg accggggcgt ggtggggcag gggcgtcttg 900	
40	ggaaatgcaa ggggtgtgca gtaaatgag agtgccttc ctcactagg tccacatctg 960	
	caggctctgg gatggagccc ccgcacaggga cccgcccctc tctgccacag acttctctc 1020	
	ccccctcccc ctggaacaca gagtggcagt ttccacaagc actaagcatt ccttccca 1080	
	aaagcccacg attggcacc ctggacattt gccccacagc cctgggaatt cactgacta 1140	
	cgcacatcat gtacacactc ccgtccacga ccgaccccc ctgtttttat taaatagta 1200	
45	caagcagagg aactccctgc taaaatgtcc tttaacaaac tgggttaaca aacgggtcca 1260	
	tccgcacggt ggacagttcc tcaacatgaa gaggaaatct cegtttataa agcctcgagg 1320	
	cattctcaag gaattacgtt gagtcaaaac tgcacctccc atgggatacg taagcaaatc 1380	
	gctcaaaaag aaagaatttc acccattggc aggggagtgg ttgggggggt taaggacggt 1440	
	ggggggcgga gctggggggt actgcacgca ccttttacta aagccagttt cctggttctg 1500	
50	atggtatttg ctcaagtatg ggaagactaac cataggggag tggggatagg ggaacccgga 1560	
	ggctgtgcga tcttgccat gcccgagttg cctggcagag ataatgtctc aggaatgag 1620	
	acgtctgatg tccccctaac ctgtggagac aaccccgccc gccccaggcg ctttgagagt 1680	
	gtgatctcgc tgggaccctt gaggctctgg atcctctggg actactcgca gcccagaaa 1740	
	gtaatccagg ggttctggga agaggcgggc aggagggtga gaggggggca gccctaggac 1800	
55	gatgaggaga gctcagtcga ggcctgaaa ggaaggagg cctcgagccc aggcctgcaa 1860	
	gcgcctccag aagctggaaa aagcggggaa gggaccctcc accgagcctg cagcaggaag 1920	
	gcacgctggg cctctagccc accagggccc atgtgtgacc tccggctccc gtgacatagg 1980	
	agggcactcg cgtcgccctt ctagcatgaa gtgtgtgggg atttgagaaa gcaacaggaa 2040	
	accatcgac tgtgaattca ggaattattc aaaaacaaag ttctacagaaa accccaagga 2100	
	cagggctgaa gtcctccgg gcaaggcgag ggcagcgac agtgatttta tttagctatt 2160	
60	ttattttatt taactacttt ctgagacaga gttatgtctt tgttgccag gctggagtgc 2220	
	acggcctgaa tcttggtcca ctgcaacctc cgtctcagcc gttccaagcc ttctgtgacc 2280	
	tccagctccc aagttagctg gatttcagcg gtgcaccacc acaccggct aattttgat 2340	
	ttttagtata gatgggcttt caccatgttg gtcaagctga tctcaaaatc ctgacactcg 2400	
	gtgatccgcc caccctcagc tcccaaatgt ctgggagtac aggcatagag cactgcaact 2460	
65	ggcctattta accattttaa aacttccctg ggctcaagtc accaccaact gtaaggagtt 2520	
	catgagattc aatttcccc ttactcagga gtatcccctc ttgatatttt tctgtaattc 2580	
	ctctgagact ggggatacac cgttctctga catattcaga gttccaagcc cactgtgacc 2640	
	tcccatggga cccactgag gggcagctgg gaggctcgag gcttcaggte ccagtgaggt 2700	
	tgccactcgc cagttagaac ctgatgtaga atcaggcgcc atttgtagaa gcaactcctga 2760	
70	actctcaatt ctcagttgtt gctgaacatc gtagaataa aagtccatcc cctcactctc 2820	
	actgggatgt agccccttcc ctatcccccc ccaggggcag aggaagtctc cctcactctg 2880	
	tggaggaagg aatgatactt tgttattttt cactcgtggt actgaatcca cgtgttcatc 2940	
	tgttggtttg tctgtttttg tcttcttctg catattcacc gttctgcca ggtctgaggt 3000	
	agtgcaatgg cgtgactctg gcttactgca gcctctcgctt cccaggttctc agtatcttc 3060	
75	ctgcttccgc cctccatttg gctgggatta caggcacccc ccacatcgcc cagctaatc 3120	
	tttgattttt tagtagagac ggggggtggg ggggttcccc atgttgcca agtctgtctc 3180	
	gaacttctga cctcagatga tccacttgcc tctgctctct aaagtctggt gattacaggt 3240	

09582246.092100

	gtgagccacc	gcgcgcagct	cagaatttta	tctgttttaga	acaactctggg	tttgtaggtag	3300
	gaagctcaac	ccactcaagt	gtttgttggt	ttttaaagcca	tatgataaatt	ttttttatgt	3360
	ttgttagaaca	ctctttgatgt	tttacaactgt	gatgactaagg	acatcatcagc	cttttcaagag	3420
	acacatacaa	tgcacatact	atactgggtt	gtctttctggg	tatacagaat	ctttcttgaa	3480
	tgcgcggagg	gtcttctctgc	caacttgacat	gggttgattat	atctcagcat	aatctctctgc	3540
	tccattctct	ttctttaaatt	ttcttcaatc	tttgagagag	tttttcaaac	gccccttctgc	3600
	aaccaagttaa	agcttaacata	taacttttga	tggaaacaaat	ttttccaaac	gcgccttctgc	3660
	ctcagtggcca	gagacaatttc	acaaacacac	ccctttaaaa	agactttaggg	atcatcaatag	3720
	ggattttctag	aagagcagcc	tatttactcta	agatttttaca	agagcgcttgc	aacctccagc	3780
	gagcgtgcaaa	gccacggagc	gggtcgaggc	ctgtttcaaat	ctgatagctca	taataataagc	3840
	aatctctctc	ggcagttctc	gagtagtaaga	aaagttacat	ttaaggtttgc	gtttcttgagc	3900
	attctaggtt	ttgcgcactc	cagcttacaag	atctcctgga	gcctctggaga	acacccagaag	3960
	tttctcgcgc	ctcttagatcc	aaaacttgagc	aacccttgagt	cttgattctct	gggaagctct	4020
	cagctgtctct	gggttggaagt	cgggggcccca	tcgtctggagg	ggaccagtctg	ctgtcttgctt	4080
	tctactatct	ggcttggaagt	cgggcctctc	acgtctcggag	tcgaaggtct	ggagcagcat	4140
	ctgcggagacc	cgagagctgcc	ctccaccctgt	tccggctcagg	atgtgacagc	atgttgacctc	4200
	ctgtctcagc	ggggccgcagg	ggggccgcagg	ggggccgcagg	ggggccgcagg	ggggccgcagg	4260
	gcgcagagct	gcgcagagct	gcgcagagct	gcgcagagct	gcgcagagct	gcgcagagct	4320
	cccggtgggtt	gattaaacaga	tttggggttg	tttgtctatg	gtggggaccc	ctctgcgcct	4380
	gagaaacctgc	aaagagaaat	gacggggctg	tgtcaaggag	cccaagctgt	ggggagaaatt	4440
	tgcaggagag	caactccggag	gtgtcccgct	gcgcgtccag	ggagcaatgc	gtctctcggtt	4500
	tgcgtcccaag	ccgcgtctcac	ccgcgtctctg	ccgcgtctctg	ccgcgtctctg	ccgcgtctctg	4560
	cggagagcgg	cggagagcgg	cggagagcgg	cggagagcgg	cggagagcgg	cggagagcgg	4620
	gcgcgcacaa	gaggtcgcac	gcacactctc	ccgatctctt	ccaatcatct	gcctctctct	4680
	cgggtttacc	caacagctctc	gcgcagtctg	ctctctctct	ctgtggccctc	ctgtagcgctc	4740
	ctctcacact	gggagacgcga	gcgcgcgcgc	ggcggggcga	cggcgcccaag	accctccggt	4800
	ccgcgccgag	cagctgtcgct	gtcgccgcgc	ggccgcggctc	ccagctgagt	ccgcgggcaca	4860
	gagccgcgag	ccgcgtctct	gtctcttgcc	gtcgccgcgc	ggccgcgcgc	ggccgcgcgc	4920
	gggttccccc	ccacagcccc	tccggccctc	ccacagcccc	ccctctctct	tcgcggccca	5040
	gcgcctcttc	tcgcggccgc	aggttcaagc	agcgtcgct	ctctctcgct	acgtggggaga	5100
	gcgcctcttc	gcgcaccccc	gcgatg				5126
	<210> 2						
	<211> 4042						
	<212> DNA						
	<213> Homo sapiens						
	<400> 2						
	gttcaaggca	gcgcctgcgc	ctgtctgcga	ctgtgggaagc	ctctggcccc	gccacccccg	60
	cgatgcgcgc	cgctccccgc	tgcgcgcagc	tcgcctctgc	ctgtcgcagc	cactacacgc	120
	aggtctgcgc	ctgtgcgcac	tctgtcgcgc	gccttgggggc	ccagggtctgg	cggttggtctg	180
	agcggctgcca	ccccgggggc	ttctgcgcgc	tggttggcca	gtgctctgag	tgcgttgcttc	240
	ggggagccag	gtgtctgcag	aggtctgcag	agcgcgcgcg	ggaggaacgtg	ctgtgcctctc	300
	gtgtgtgcgc	gtgtgtgcag	gccgcgcggg	cccccccccg	ggccttccac	accacgcgtc	360
	cgagctatgc	gcccaacacg	gtgacgcagc	caactcgggg	gagcggggcg	ttggggctctg	420
	ctgtgcgcgc	ctgtggcgac	gaagtctgtg	tttcaactgt	gcgcctctgc	gcgctctttt	480
	gtgtgttgcc	ccccagctgc	ctctaccagg	tgttcggggc	gtgtctgagc	gtgtctgagc	540
	ctgtgactca	gaactcatgc	cttcaggagg	ccgggtgtcc	cttggggctc	ccagccccgc	600
	aaaggcgctg	gcgcgggggc	agttcgcagc	gaagttctgc	gtgttccaaag	aggccccagc	660
	gtgcgcagtg	cccttagggc	gagcgcagcc	ccgltggggc	ggggttccag	gcccaacgct	720
	gcggagcctg	tggacagctg	gagcgttgtt	tgtgttgatt	gtcactctgc	agacacccgc	780
	gcggagccac	gagcgcagc	gtgtgcagct	gtggagagag	gcacccctgc	gcacccctgc	840
	gcgcgcagca	ccacgcgggc	cccccatcac	catcgcgcgc	accacctgc	ttggagacac	900
	cttgtcccc	gggtgttcgc	gagacacatc	acttctctat	ctctcaagc	gacaagagac	960
	agctgcgcgc	ctctctctca	ctcaagcttc	tgaagccagc	ctcgaatgac	ctctggaggc	1020
	tcgtggagac	catcttcttc	gtttccaagc	ctctggatgc	aggagatccc	ccagagttct	10

	agatattcca	tgttctataga	ttgggaagat	aaatactgtt	aaatgttcca	2580
	agcaatttac	aaattccaatg	caactccctat	taaaatacta	atagcgttct	2640
	agaagaaca	atttcaaatg	ttgtacacgaa	ccacaaaaga	cccgagaatg	2700
5	agctgaccaca	aagacaacaaa	ctgggaagcat	ccacttacc	gacttcaaat	2760
	agctatagta	accacaacta	catgtgtactg	gcataaaaa	agatgagaca	2820
	gaacagaata	gagaaactcag	acaaataact	atgcattctac	agtgtacata	2880
	agggtgcgaag	aacatctatt	gggggaaga	taactctctt	tttttgacaa	2880
	ctgggatacc	atatgcaaaa	taacaaact	agaactctgt	gtcggagaaa	2940
	aaatcaaaat	ggatgaaag	cttaaatcta	aaactctaaa	ctctccact	3000
	aaacacggag	aaactctcca	ggacatttga	gtgggcaag	actttcttag	3120
10	caggacacag	caaccaaaag	aaaaacagac	aaatgggat	atatacaagt	3180
	tgccacagcaa	aggaacaacat	caaccaagag	agaagacaga	ccacagaagt	3240
	ttgttcaaat	attctatcaa	caaggaaatta	ataaccagta	tatataagga	3300
	ctctataaga	aaacacacta	tttcaaaat	agaagcaaga	ctcgggtaga	3360
15	cattttctca	aaataagtc	acaaattgca	aacaggcat	tgaataattg	3420
	ctgatcatca	gagaaatgca	attcaaaact	actatgagag	atcatctcat	3480
	atcgctttta	ttcaaaaagac	aggcaataac	aaatgccagt	gaggatgtgg	3540
	acctttggac	actgttgggt	ggaaatgaaa	ttgtccacc	tatggagaac	3600
	ttctctaaaa	aactaaaaat	aaggtaccca	tacagcaatc	ccattgtctg	3660
20	caaaaaaggg	aatcagtgta	tcaacaagct	atctccact	ccacatttcc	3720
	ttctatagcag	ccaaggttg	gaagcaacct	cagtgctcat	caacagaca	3780
	aaaaatgtgt	gcacatacac	aatggagact	tacgacgca	taaaaagaaa	3840
	tcagttgcaa	cagcatgggg	ggcactgttc	agtattgtta	gtgaataaag	3900
	agaacacaaa	ttttcatgtt	ctcccttact	ttgtgggagca	aaatataaaa	3960
25	agaaataag	gagaaatgtg	gttctagagg	gttggggagg	aggggagca	4020
	tatgttatgt	tatgtttatg	aatcaactga	agagtataat	gggtttgttt	4080
	gaagaggtata	atgtcttgaag	gtgacagata	ccccatttcc	cttgatgtga	4140
	tttatgtctt	gtatcaaaat	atctcatgta	tgtatatagt	ataaacctta	4200
	aattaaaaat	tttgatggcca	cttcatgtcg	taatcccgac	actttggggg	4260
30	gcggagggcgg	gttgatccac	tgaggtcagg	agtttgaac	cagttctggc	4320
	aaacctgtct	ctactaaaga	tacaaaagat	agccaggggt	ggagacacat	4380
	ccaactctc	agggagctga	gacaggagaa	ttgtcttgaa	ctggagagca	4440
	tgagcgcaga	tcatgccact	gcactgcagc	ctgggtgaca	gagcaagact	4500
35	acaaaaacaa	aaaaaagaag	attaaaaatg	taatttttat	gtaccgtata	4560
	ctactatatt	agaagttaaa	aattaaaaaa	attataaaag	gttaataaac	4620
	aaataaagaac	aatgtatgtg	gggtttctag	ctttcgaagt	agtaaaagt	4680
	tggcagaat	ctggggaggg	aacagtgggt	gttactgtgt	ttagagctgt	4740
	taagtgaact	tttttcaact	aaagacagg	tgggaagaat	aaagagacca	4800
40	cttaaaaaaa	ctgttaataa	tgttgaaagg	taattcttat	taattaccaa	4860
	tatctctaaa	actagagctg	agaattggca	cgctgtatca	cacgtctctc	4920
	tgtctttttt	cttgtgtgct	tggagatttt	cgatgtgtgt	ttcgtgtttt	4980
	actctgtaaa	actctgaaac	gaaaaatagg	gggtgatttc	tcagagaaa	5040
	tggcaggaag	caggtggctc	tgtggaaact	agccacttca	actttcaagg	5100
45	agaacccagg	tgaacggcag	aggcctgatg	accggagacg	aggaaaagtc	5160
	gggcgctgag	agaagctgct	cggttggtgag	cagcgcatga	agtgccctt	5220
	gcaaaagtat	ctctggatac	catctggaaa	aggcgccagc	ccgggaatga	5280
	agcctccctg	tcaaaaccag	gccagcagct	atggcgccca	cggggaggtg	5340
	agaagagata	aggcactctg	aaagttaggt	taaatctttt	tttaccctga	5400
50	agggtgtatt	cttgagggaag	cttaggttag	gtgccttttt	taaaacagaa	5460
	gcacctttct	caagggaana	ccagacgccc	gtctctcggt	catctaactc	5520
	ccctctcttg	ccctgcgggt	ttctgatcgg	tcaacagatga	ccccctggga	5580
	gcccgctgtg	aggaccctct	tgcaaaaggc	ggcagacacc	cccgccctgt	5640
	ctgagcctgg	cttaataaca	aaatgggatg	tgtcgggggg	cgagacagca	5700
55	aaagagctta	actctatgag	taaaatcaac	cttgatacatt	ctggaaatgat	5760
	ttctaatatt	ttcttaaat	ttcttaaaat	accatcagga	cttcagaaat	5820
	aaacacggaa	ctgagcttatg	tttgcacaa	tcaaaagact	taataaacat	5880
	atttttgcgc	actgaactat	ttttatgtgt	ttcataaggt	ggcttagggt	5940
	gtacacagag	agaaggctctg	ggcgaggggc	tatgagacag	gcagggccac	6000
60	ttcccccggc	tgggagagctg	acacgacgac	cttcacacgt	ctccctagg	6060
	tgggcacacg	cgagggggc	cagctgtggt	gtgactcagg	gacctctagg	6120
	ggccacccact	actaaaccag	gaagtacagg	agctcttgaa	ccgtggaaac	6180
	ctgctccgcc	tgtctccctg	gggtgggtcaa	gggttaatgaa	gtggtttgtca	6240
	atgtaaatcta	cacgactctg	atgcatggga	ccgttctctc	catcatattt	6300
	cccaaggact	gaatgatctc	agcaacttct	tcgggtgtga	caagccatga	6360
65	tacaaacacac	actcttttacc	taggccacaa	gagacagcac	cacacacctc	6420
	agtcacagg	agatgaagct	gctttacgc	agcagagctg	ggtagacata	6480
	agtcgtctcc	tctagactag	tagacctctg	caggactctc	ccagagattc	6540
	tgtctgtctc	cgaggggccc	actctgccc	gagactcagc	ctgggggtgc	6600
	cagcctctgc	tccacacctc	ccgctccag	gctccagctt	cttccagacc	6660
70	ctgggtgggg	cggtgttccag	cgctactgtg	tcactcgtcc	ccactgtgct	6720
	acgtagtaac	cacggtttac	agctcacatg	gtgtgtgtgt	ttcttcccca	6780
	ctgttgaaag	gagagatttc	tggcctctcc	agactgggtc	cttgtagctt	6840
	gttggggccc	gatgcaggtt	ectggcgtcc	ggcttgcaac	tgactcccat	6900
	tcccctctct	ctgtcatctg	ccggggctcc	ccgttgtgtt	ttctgttttc	6960
75	acagggctgc	gctgcgtgtg	tctctcccgc	ctagggtctc	ggggtttttt	7020
	acagggggcat	gtggggccag	ggcgctctct	ggaaaatgcaa	taatgggtg	7080
	agtgcccttc	ctcactcagg	tccacgggc	tcacagggc	gagcagggc	7140
	cccgcctctc	tctgcccacg	actttctcgc	ccccctctc	ctggaaacaa	7200

09599694262100

	ttcccaaacg	actaagcatt	ctcttcccaa	aagaccacgc	atgggacccc	ctggacattt	7260
	cccccaacgc	ctctgggaatt	cacgtgacta	gcacatcat	gtacacactc	cgctccacga	7320
	ccgacccccg	ctgtttttatt	ttaatagcta	caaaagcagg	aaatccctgc	taaaatgtcc	7380
5	tttaacaaac	tggtttaaaca	aacgggtcca	ctcgacaggt	ggacagttcc	tcacagtgtc	7440
	gaggagagct	cggtttataaa	agcctgcagg	catctcaagg	gaattagctt	gagtcacaaa	7500
	tgccactctc	atgggatacgt	tacgcacaat	ctgcataaag	aaagaatttc	accctcagct	7560
	agggagagtg	ttaggggggg	taagggacgt	gggggcggca	gctgggggct	gctgcacaga	7620
	ccctttacta	aagcagattt	ctctggtctg	atggatattg	ctcgaatttg	ggagactaac	7680
	catagggggg	tgggtgattgg	ggaacccgga	ggctgtgcca	cttttgccat	accgacagtg	7740
10	ctctggggcg	ataatgtctc	agagatgcgc	agctcctgat	tcgcccaaac	ctgtggagac	7800
	aaccgcccg	cgcccaaggc	ctttgcaggt	tgatctccg	tgaaggacct	gaggtctcgg	7860
	atctctcggg	actacctgca	ggccccgaaa	gtaatccgg	gggtctggga	agaggccggc	7920
	agaggaggtc	gagggggggc	gcctcagagc	gtagggagca	gtcagttcga	ggctgaanaa	7980
	ggaggagggg	ctcgagagcc	agggctgcaa	gcgcctccag	aagctggaaa	agcgagggaa	8040
15	ggagaccctc	acggagcctg	caggcagaa	gcacggctgg	cccttagccc	acagggccgc	8100
	atctgtgacc	tcgggctctc	gtgcataagg	agggcactcg	cgctgccctt	ctbagatgaa	8160
	tggtgtgggg	atttgacaaa	gcacacagaa	accatgcac	tgtgaatcta	ggattatttc	8220
	aaaaacaaag	ttacagaaa	catccaaagg	cagggtcccg	gtgcctccgg	cgaggtccgc	8280
20	ggcagagcag	agtgtattta	tttagctatt	ttattttatt	tacttacttt	ctgagacaga	8340
	gttatgtctc	tggtgccag	gctggagtgc	agcgacatga	ctttgctcta	ctgcacactc	8400
	cgctctcctg	gttcaagcaa	ttctcgtgcc	tcagctctcc	aagtactctg	gatttcagcg	8460
	gtgcacaccc	acaccccgct	aatttgttat	ttttagtaga	gatgggcttt	caccattgtg	8520
	gtcaagctgc	ttcctaaatc	ctgacctcag	gtgatccgct	cacctcagcc	ctccaaagtg	8580
25	ctgggttatc	agggcatgag	cactgcacct	ggcctattta	accattttaa	aacttccctg	8640
	ggctcaagtc	acacccactc	gttagggagt	catggagttc	atcttctcct	ctactcagga	8700
	gttactcttc	gtctgatttt	tccttagact	ttcttgagct	ggggtacacat	cgctctctga	8760
	catattccaa	gtttctgtga	ccactgttta	tcctatggga	ccactgcag	gggcagctcg	8820
	gaggctgtgc	ctctcaggtc	ccagtggggt	tgccatctgc	cagtgaatac	ctgctatgaa	8880
30	atcagggcgc	aagtgtggac	actgtcctga	atctcaatgt	ctcagttgtt	gctgaatacat	8940
	ttagaataata	agattccatcc	ctcctacttc	actcgggttg	agccctctcc	ctatcccccc	9000
	cgaggggggc	aggagttctc	ctcactctctg	tggagggaag	aatgacattc	tgttattttt	9060
	cactgtcgtg	actgaatcca	ctgtttctat	tgttggtttg	tttgttttgt	tttgagagcg	9120
	tggttcaact	ttgttgcctc	ggctggaggg	agtgcaatgg	cgcgactctg	gcttacttca	9180
35	cgctctcgct	ccccagttca	agtgtatttc	ctgcttcgcg	ctcccatttg	gctgggataa	9240
	caggcacccg	ccacatgcc	cagctaattt	tttgtatttt	tagtagagac	gggggtgggt	9300
	gggggtccac	atgttgggca	ggcttggtctc	gaactcttga	ctcagataga	tcacactctg	9360
	tgtctctctc	aaagtgcctg	gattacaggt	ctgcagacac	atgcagctct	cagatattac	9420
	tcgtttttaga	tgctgagtag	tcactcagct	gaactcagct	gtttgttggt	tttgtgtgtg	9480
40	tttaagccaa	tgatagaatt	tttttattgt	tgtagaaca	ctctgtgatg	tttacaactg	9540
	gatgactaag	acatcatcag	cttttcaaa	acacactaac	tgcacccata	atactggggg	9600
	gtctctctgg	tatcagcaat	cttcatlgaa	tgcggggagg	tggtttctctg	ccatgcacat	9660
	gggtgttaatt	actccagcat	atactctctc	ctcttctctc	cttttttaaa	acttttaaa	9720
	tgtgtttctc	atgttggttt	ctctgcagag	agcacagtta	agctacaact	taacttttgt	9780
	tggaacaaa	tttccaaacc	gcctctttgc	ctcagtgcca	gagacaaatt	acaaacacag	9840
45	ccctttaaaa	aggcttaggg	atcactaagg	ggattttctg	aagagcacc	tgtatactca	9900
	agttatttca	agacgaggtc	aacctccagc	gagcgtgaca	ccccaggagg	ggtgcgagcg	9960
	ctgttcaaat	gctagctcca	taataaaagc	aattctctcc	ggcagtttct	gaaagttaga	10020
	aaggttacaat	ctaaggtgtc	gtttgtttag	atttcagctc	ttgcgacctc	cagctataga	10080
	atctctgcga	gctctcggtg	gaccagaaag	ttttctcgcc	ctcttagact	aaactcagc	10140
50	aaccggaggt	ctggtatcct	gggaagtctc	cagctgtctc	cggtgtgtgc	cggggcccca	10200
	ggctgtggag	ggaccagttg	cgctgtggct	ctcactgcct	gggtggaagt	cgggcctcct	10260
	agctctgcag	tcgagagctt	ggagccaggt	gcctgcagcg	cgaggtcgct	ctccacctgc	10320
	tgcggggcgg	atgtgacacg	atgttggcct	catctgcag	acagagatgc	ggggccccag	10380
	ctcaggagcg	tgtgtgctg	tgtgagagcg	ccggtgcgag	gcagacagca	gcgcctggct	10440
55	ccatttccca	ccctttctcg	acgggacgcg	ccggttggtg	gataaacaga	tttgggggtg	10500
	ttgtctcatg	tgtggggacc	ctcgcgcctc	gagaacctgc	aaagagaaat	gacggcgtct	10560
	tgtcaaggag	cccaagtctc	gggggaagtgt	tgcagggagg	cactccggga	ggtcccgctg	10620
	gcctctccag	ggcgaacatg	gtcctcgggt	tcgtccccag	ccgcgtctac	gcgcctcctg	10680
	ccctcccttc	agctcggcca	ttcagtggtc	cgagagcccg	acgcgcccg	tcggaacctg	10740
60	gagcagccc	tgggtctccg	gtctcagcca	ggcgcccaag	ggtcgcccga	cgacactgtt	10800
	cccgaggcct	ccaatcatg	gcctctctc	cggtttacc	caacagctag	gcgcatttga	10860
	ctctctctc	ctggggccct	cgctggcgct	ctctgacctc	ggagcgcgga	cgggcgcgcg	10920
	ggcgggggag	cgcgggccag	acccccgggt	ccgccccgag	cagctgcgct	cgcgggccca	10980
	ggcgggggct	ccagtggatt	cgcgggcaca	gaagccccag	cagctgcctc	ctcagttggg	11040
65	gaggagagcg	ggacccggcg	accctgcctg	ccctctcaac	ttccagctcc	gctcctcccg	11100
	cgcggaaccc	ggccctgccc	gacctctcc	ggcccccggg	ccagctcccg	tcggggccct	11160
	cccgagcctc	ccctctctc	tcggcgccc	cgccctctcc	tcggcgcgag	agtttcagcg	11220
	agcgtctcgt	ctcgtctcgc	acgtggggag	ccctggcccc	ggccaccccc	gcgatg	11276

<210> 4
 <211> 104
 <212> DNA
 <213> Homo sapiens

<400> 4
 gtggggcttc cgggggtcgg cgtccggctg ggggttaggg cggcggggg gaacacagca 60
 catgcggaga gcagcgcagg cgactcaggg cgcttccccc cgag 104

<210> 5
 <211> 8616
 <212> DNA
 <213> Homo sapiens
 5
 <400> 5
 gtgaggaggt ggtggccgtc gaggggccag gccccagagc tgaatgcagt aggggctcag 60
 aaaaagggggc aggcagagacc cgtgtccctc tgtctccatc gcaacgtggg cacaacgtggc 120
 ttctcgctca ggcgtctcag tgcacaacgt gatctctgcc tctgtctccc ctctgtgttg 180
 10 tctgtgtataa acttacgagg ttccacttca cgttttgatg tgacacggctg ttccagcgtc 240
 cgaggccaga cagtgtaaca gaggaggtcg gcgcggcgac ggagccggg tgcgcggcaa 300
 tggggggaag tgtctggaag cacagacgtc ctggcgaggg tgccctgcagg ttacctataa 360
 tctctctcgc aatttcaagg gtgggaatga gaggtgggga cagaaacccc ctctctctgg 420
 15 ggggtggagg taagggtttt gcagggtgac gtgggtcagc aatatgcagg ttgtgtgtta 480
 aggttaattt gtgtctgac ggcacaggtc ggtagctcac cgaggtatca ccaactgatt 540
 ggggaactga ggcagggtga tcaactgagg tcaggagttt gagaccagcc tgaccaaat 600
 ggtgaaaccc tatctgtact aaaaatacaa aaattagctg ggcattggtg tgtgtgcctg 660
 taatccacgc tacttgggag gctgaggcag gagaatcaat tgaacccagg aggcggaggc 720
 20 tgcagtgaac tgagattgtg ccattgtact ccagcctcgt cgacaagagt gaaactatg 780
 ttcttaaaaa aaaaagcgtt cgttgatgtg gccaggacag ggtgaaggga gggagataag 840
 actgtctccc agcaacgalt ctgtgccat cttaagtat gaagagggcc acatggggagc 900
 agaggacagc agatggctcc acctgtcgag gaagggaacg tgtttgtggg tgttccaggg 960
 25 atgtgtctgc tgggcccgtc cgtgtcccca cectgtttt gcgtatttga tgttgaggaa 1020
 cctccgtctc agcccccttt ttgtcccaaa tgctccaggg cctacagctg gcagctagaa 1080
 gaagtccgga tttaaccccc tcccaacaaa ctcccaagg acataagact tctgtgtatc agtgcagat 1140
 cagacaaggga ggggtgacct ctgggggtcc tttttttctc tttttttctc tatgggtggt 1200
 aaaaagtata taactgaga ttgtgacctc taacaacctg tctgtgtatc agtgcagat 1260
 30 tgcatactgc cgggtgttta cagcagggtg cttgaaatgc tgcgtcttgc gtaactgaa 1320
 gtccctaccc atcgaaacgc agctgcctca cactgtctgc ggctcagtg gtaacacgct 1380
 agtcaacaaa cgtctatgca acccagtttt gctttttgtg tccagctctc cttcgtttag 1440
 gagagtttga gtctctgatg caggactctg cctgtcatgt ctgtctctgt acctcaagt 1500
 aggtcacaat ctgcccctgt ctatgacagg gagtgaggcg tggctcccgg gtgcccctgt 1560
 35 caagtgcaag gtgagtgagg cgtgtccccc aggtgtcccc gtcaagtgta ggtgagtgta 1620
 gcgcggccgc ccgggtgtctc ctgtccccgt cagcgttatg gaggtgtgac ccccggtgt 1680
 cccgtgcacg tttagggtga gtgaggcgcc atccccgggt gtcccgtgca cgttagaggt 1740
 gaggtagggc tggctccccg gtgtccccgt cccgtgcagg gtgagtgagg ctatgtcccc 1800
 ggtgtccccg tgcacgtgca gggtagtgga ggccgcgtcc ccgggtgtcc cctcagtgag 1860
 40 taggttagtg gaggcgagc agtgagggtg agtgagggtg tggtagagtg tgggtgggtg 1920
 gtccctgggt gttccctcca ggataagggt gaggtaggca cgtccccggg gtgtccccgt 1980
 ccagtgcaag ttaggtgagg ccgcggcccc ggggtgtccc ctcaagtgca gggtagtgta 2040
 ggcgtgtccc ctgggtgtcc ctgtctcgtg tagggtagtg gaggctctgt cccccaggt 2100
 45 ccttggcgtt tgcctacttg agcttgctcc tgaattgttg cttcttctat agccacagct 2160
 ggccgggttg cccattgctt gggtagagtg tgcaggcgca gtgctgtccc ccaagcctat 2220
 cttcttgatg gttcggtctc tcttggtcac ctctccgttc cactttgata cggggacaag 2280
 ggaactcagg tctcgcctc ccgctggcca ggaactgcag ccaacagctc gtaactgcctt 2340
 50 gctctgtgtg ggtccggctt gctcacacag tgcgccacac atgcatgtgt ccaatactcc 2400
 tctccagact ttctctatgc cagggtgga ccttgggtgt cctgtgtctg gtcgcaagct 2460
 ttgcttgaga catcccaaga aggtttctct gtgccctgaa ggaagacaag tcaacccaag 2520
 55 gctccacttt gtcctgtttt ctcccaagct tgcgaagctc tggcgagctg tgggtgggtg 2580
 gcaacgtgtg ccaacttatt ctgggaacct ccgcctcact gcttaggctg ggtctcagct 2640
 ccagctgcgc cctcacatgg attgaagctc agccaacagg tggagtgtct ctgtctctct 2700
 cctgctctga gaccacaagt gagggcgggt. gctcccgca cctctgcga cactctgca gaactccctc 2760
 60 ttgggtctta gttttgaatt tcaactgatt acctctgac ttctctatcc tcaactgtct 2820
 gcttttctct ggtttatctt ttcatctctt ttetagtctc ttagttagt atcgtcttct 2880
 cctctaaagt gctcggtctc agtcactgtg tggttacag acctgtccat ttgaaatttg 2940
 65 cacttcaag ttgtcttaaa atactcaaa gtgttaaac tttttttaa tattcttatt 3000
 ctgtgatttt ttttttgtg cagcgtgtgt tttgacgtga aatcatttt atactcagta 3060
 cttttaagta ttctttagct tattctgtga tttctttgag cagttagtta ttgaacact 3120
 gttttatgtc aagatacgta gagtacaa acacgtagag tatcttaagt tatcatttta 3180
 70 ttatgatgtt ctaactcagt tgtgtagtggt tctgtataat accaattatt tgaagtttgc 3240
 gggagctggt agtgcgtctc atgtcgtgtc actatctcca gcttaggctg ggttaatttg 3300
 acactctgtc aatagtgggc atgactgttc actatctcca gcttataag tgcacgtgct 3360
 aagcttctgt cctcctctag atgcatgaaa ttccaagaag gaggccatag tccctcaact 3420
 75 gggggatggg tctgttcatt tcttctgagt ttgttagct tatgtgagc attttaggt 3480
 gcatgacagt ggtagaattt ttatccttgt gatgtgaa ttttttggag accttatgt 3540
 ccttagtaat ctgataattc tttttttaaa ttgctcttag taccgcaca cggggcttct 3600
 ttgtagtag atattctctc agtggtgtt ttgtgcttt aatttataa tatataatc 3660
 ttttttttt ttttgagaca gagtcttggc ctgtccgcca ggttgagtg agtggtgtga 3720
 80 tccacagtgca gtgtaacttt taacctctgg cctgagcgtc cctctcactc cagctctctg 3780
 agtagctgga actcgagaca cgcaccgctc cacttggtta atttttaatt ttttttggg 3840
 gcagggtgtc tctgtgtgtg cccagcgttg tctcaaacct tggactcaca gggatccatc 3900
 taactcggct tcccaaaagt ctgaattaca ggcataagcc acctatctgt gcttaatttt 3960
 85 caactctat atatcttat agtggtgga tgcctcttta acagagtgga ggtgaatttc 4020
 caatcagtc tgaacagctg tgtttaaact gataactgta tttattttta ttttttttgc 4080
 actagagacc cgcctgtgtc actctgactc tccacttggc tttgtcatgt cctgtctccc 4140
 90 ttgtttctca ccaacccttg ggttgcatct tgcgtttccc gccagagtggt tgttgatcct 4200
 ctgctgtcct cctggtcact gggcatttgc ttttatttct ccttgcttag tgttaccccc 4260
 tgactttttt atgtctgttg tttgcttttg tttatttga cagctcactc cgttaccaca 4320

	ggctgcgaagt	aaatggcaca	atctgcagat	actgcacagt	ctgcctctcc	gggttcacaa	4380
	gtttcattcc	ctcaacacca	ttagtagctg	ggatataccg	gcgcccaacc	accgctgcgc	4400
	taatttttgc	atattttatg	gagatagctg	cttcacattg	tggccagagc	tgctctcaac	4500
5	ctctgcaccc	aagttagctg	gccgccttgc	ctccccaacg	tgtcggagat	acagggtgca	4560
	gcacccattc	ccggcgtata	tatgtttttt	aaaatagatt	ctgcnaaacg	ctgaccctta	4620
	tcctgagcaa	tgaagccctt	atgtagtatt	ctgctgcttc	ctcccccacg	ctgacatttc	4640
	ctctctccgc	ctctctcttc	gagctctctg	catcttgaca	ccccccacac	ctaaagctta	4700
	ttaatattgt	tttcctgttt	gaagtgtttt	gtagactttc	cccccgcctg	ctttttctcc	4800
	tttgttcccc	ttctgtcttc	ttgtccagcg	ccgcgctctg	gtgttccctt	ctttgtctct	4860
10	tgcggtgttc	ttctgtcttc	ttatctgtctg	taaaccccag	cttttaacct	gtgtgtctcc	4920
	atggcattca	gcgacgagtc	gggaccttgc	ctctatagat	acagatgaag	atgtggggac	4980
	tcacagagag	ccgggttatc	tggcgcgtgc	agctctcttc	gcagctcttc	taaccctcag	5040
	ctacagcagt	ctctgtcttc	ccgcgcttgc	gaggttgaaa	ctgcgcnaac	ccctgcggtg	5100
15	catgtgcgtg	tcctgtgttc	ccgcgcttgc	gttgttgaaa	ctgcgcnaac	ctgcggtgtg	5160
	gcgccacatc	ctgcgtgttc	gctctgcggg	ggagttgttc	ctcttctcct	cttgttgttg	5220
	acacagacaa	aagattagag	ctctcagagc	gtgtgcctca	accaggacat	gaactgagcc	5280
	atbtgtgata	ttttaaaaat	ctctagctgtc	gcgcgctggc	ctgcagcgtt	atactccagca	5340
	ctctggagaa	ccctcagcgg	tgagatacga	gtctcagagag	ctgcagacat	ggcggtgtgc	5400
20	catgtgaacc	ctctctcttc	acagctcagg	ctctcagcgt	ctgcagacac	ggcggtgtgc	5460
	tgtaatacca	gctactctgg	acagctgagc	aggagaattg	ctgttaacct	ggagtttgaa	5520
	tgctcagtag	gcgcacatgt	ccccctgcga	ctctcagcgt	cccaacacag	gaagactctt	5580
	ctcaaaaaaa	aaaaaaaaaa	aaaaaaataa	aattctagta	gcccaattaa	aaaaataaaa	5640
	atgaaaaagt	gaatttaagt	taataataag	ttttaactga	gcgcagcagt	ttccacaacta	5700
25	atcattttat	gggtcttatg	gtggagagag	ctcagtgaca	ctctctgcct	ggacctcttc	5760
	ctctctctgc	ggcgaagagc	gtgttaccag	atgtgtcagc	ctccggatga	gggtccagag	5820
	gctctctcat	gctgtgatgt	catgttcccg	atgtgtcccg	ctctgggatg	gggtccagag	5880
	ccctcagtag	gctgtgatgt	gtgtgtcttc	gatgtgtcac	gtcagggggt	aggttctccg	5940
	ccctcgtctg	agctggatgt	atgtgagctc	gatgatgcag	ctccgcggcg	agggtccggc	6000
30	ccctcgtctg	agctggatgt	atgtgagctc	gatgatgcag	ctccgcggcg	agggtccggc	6060
	ggcctctctg	tgagctgtgt	gtgtgtgttc	tgagttgttc	agggttcctc	tgagcttacc	6120
	aggcctctgc	tgagctgtgt	gtgtgtgttc	tgagttgttc	agggttcctc	tgagcttacc	6180
	aggcctctgc	tgagctgtgt	gtgtgtgttc	tgagttgttc	agggttcctc	tgagcttacc	6240
35	aggcctctgc	tgagctgtgt	gtgtgtgttc	tgagttgttc	agggttcctc	tgagcttacc	6300
	cagacgctgc	cagacacatc	gggtgacgtg	atatgcggtg	ctccgagatg	cagagttctg	6360
	gttgaagttc	cagacgcttc	ctctgaagtc	gatgttggtg	ctccgagatg	cagagttctg	6420
	gtgtgagatc	accagccgctc	gctctgagtc	ggatgtgtgc	gtgtgtgtgc	gggtgtgcac	6480
	gtgttgaaag	agggttcgag	ccctccggtg	gaagctgatg	cgaggttcca	gatgtgtcac	6540
40	gtgttgaaag	agggttcgag	ccctccggtg	gaagctgatg	cgaggttcca	gatgtgtcac	6600
	gtcggcgagc	agggttcgag	accctccggt	gaagctgatg	ctgcggtctc	ggatgtgtcac	6660
	gggttcgagt	agggttcgca	ggccctccgt	gaagctgatg	ttatgaagtc	gtgtgtgtac	6720
	gggttcgagt	agggttcgca	ggccctccgt	gaagctgatg	ttatgaagtc	gtgtgtgtac	6780
	agggttcgag	tgagttcgcc	ggccctccgt	gtatgtgtgc	gtgtgtgtgc	gggtgtgtac	6840
45	caggttcgag	tgagttcgcc	ggccctccgt	gtatgtgtgc	gtgtgtgtgc	gggtgtgtac	6900
	caggttcgag	tgagttcgcc	ggccctccgt	gtatgtgtgc	gtgtgtgtgc	gggtgtgtac	6960
	caggttcgag	tgagttcgcc	ggccctccgt	gtatgtgtgc	gtgtgtgtgc	gggtgtgtac	7020
	caggttcgag	tgagttcgcc	ggccctccgt	gtatgtgtgc	gtgtgtgtgc	gggtgtgtac	7080
50	caggttcgag	tgagttcgcc	ggccctccgt	gtatgtgtgc	gtgtgtgtgc	gggtgtgtac	7140
	ctcagagctc	gggttgagtc	ccaggccctc	ctcgtgagtc	gatgtgcgac	ctgtcccggt	7200
	ctcagagctc	gggttgagtc	ccaggccctc	ctcgtgagtc	gatgtgcgac	ctgtcccggt	7260
	ctcagagctc	gggttgagtc	ccaggccctc	ctcgtgagtc	gatgtgcgac	ctgtcccggt	7320
	ctcagagctc	gggttgagtc	ccaggccctc	ctcgtgagtc	gatgtgcgac	ctgtcccggt	7380
55	ctcagagctc	gggttgagtc	ccaggccctc	ctcgtgagtc	gatgtgcgac	ctgtcccggt	

```
<210> 6
<211> 2089
<212> DNA
<213> Homo sapiens
```


	<400> 6	
	gtactgtatc cccacgccag gccctctgctt ctcgaagtcc tggaaacacca gccccggctc 60	
	agcatcgccc tgtctcccaat tgccctgtgct tccctggctg tgcagctctg ggcctgggagc 120	
5	cagggggcccc gtcacaggcc tggccaagt ggaattctgt ggaagctctg actgctcgga 180	
	gcctcagcttc tcttacttgt aaaaatcagg gtttctgcca agtggctctct aggggttgtta 240	
	aagcagaagg gattttaaatt agatggaaac acctaccata gccctcttgc ctttcccctg 300	
	gatgctgggtc tgattctctc tctcttcttt ttctctcttt tgaagatgag tctcaactctg 360	
	ttggccaggc tggagtgca tgccataatc ttggctcaat gcaacctcca cctctcgggt 420	
	ttaaaggatt caccagcctc agcctctctaa gttagctggga ttacaggcag ctgcaccacc 480	
10	gccctggctaa tttttgtact tttaggagag accgggtttc accatgttgg ccagggtctgt 540	
	ctcgaaacca tgacctcagg tgatccaccc accttggctt cccaaagtgc tgggtttaca 600	
	ttctaagcca ccgtgcccag cccccgattc tctttttaa ttcatgtgtc tgattgaatc 660	
	gcatctcat tggattcaag tcatgagag ataaaatccc accacttccc tgcactctgc 720	
15	caggggagac ctgtcgaggg agcaacttgg gatagagag ttccaccatg agctaacttc 780	
	ttgggtgctg catttgaatg gctgtgagat ttgtctgca atgttcggct gatgagagt 840	
	tgaagtgtg acagattcaa gctggatttg catcagtga ggaacgggag cgtggctctg 900	
	gagatgccag cctggctgag cccaggccat ggtattagct tctcgtgtc ccgcccagg 960	
	tgactgtgga gggctttagt cagaagatca gggcttcccc agcttcccgt cacactcgag 1020	
20	ttccctggggg gccctgtgag accccatgcc ccaaaatcag atgtctcgag agggagctgg 1080	
	ctgtcagagg ttacacagct ctgggcttgg gaccggcag tgggtctggg 1140	
	gcattttctc tgcactctgg ggagggctca ggttttcccc tgcgggaacca gtttaatac 1200	
	aatgcacattt acttagactt tacactgatt taattgtgtg cgacccaaca tggctcattt 1260	
	accagtattt tggaaagaat ttaattgggg tgaccggaa gacagacag acgtggttgtt 1320	
25	ccccaaagat cctcctgtgca ctactgggag ttgtgttctg cctggggggg cctggagccc 1380	
	ccctccctctt ggaacagggtta ccgtgccttt tctactctg tgggctctg cgtgcgggtc 1440	
	agggccagag ctcggagaca cccggggccc caggtctcac gaggatggca cttgacagcc 1500	
	acagatgcce aggtccaggt gtggccgctc cagccccctg gcccctatgg gtgggttttg 1560	
	gggaaagagc caagggcaga ggtgtcagga gactggttgg ctcatgagag cttactctgc 1620	
30	tctcttgctg agctgcctct agcagcctct ccccgccctc ctcatctgaa ggaattgtctt 1680	
	cttttactc tgggggtctg cctgggggca gcccttgagc accccagtgg ctgtaccaga 1740	
	gggacagcca tggctctggg agggctatgg gttcacgtg cctcagagtc agcctcgagc 1800	
	caggtctcct gtgtctgagt gtgggacagt caccctgggg gttagccggc ggaactgtctg 1860	
	tcccagaggt tgactatagg accaagtggtc caggtgccct ccaagttaga ggtgcttcag 1920	
35	aggcgtctgg ctggcatggg tggacgtggc cccgggcag gccctcagcg ttgctgtcgg 1980	
	tgggtgccct gaggcctcac tgagtcgggt ggggcttgtt ggttccccgt agcttcccc 2040	
	tagtctgtg tctggctgag caagcctcct gaggggctct ctattgcag agcttcccc 2080	
	<210> 7	
	211> 687	
40	<212> DNA	
	213> Homo sapiens	
	<400> 7	
	gtggctgtgc tttgtgttaa cttccttttt aaacagaagt gcgtttgagc cccacatttg 60	
45	gatcagcatt agatgaaggg cccgaggagg gggccacggg acacagccag ggcctatgca 120	
	cgggcacca ccaattgtgc gcaactgtag gtggccgagg tggcgtgccc tccagaaaag 180	
	cagctgtggg gtgtaggggg agctccctgg gcagggaacg gctctgagga cccacaaga 240	
	cgccggggc agggcttgga tgcagcagg ccagagctc tggattccgt tctcgtgtgt 300	
50	gtggcagcc tccgtgcctc tccgtctatg gggccgggg accagacca cactgtccag 360	
	agcccacccg gctctgagga tcttggaact tgcgccact cctctcgac caccctctg 420	
	ggctcggggt gctcggtgta cccgctcatc tgaggagagt ttgggggtgag gtggacagag 480	
	tgtgtggcatg aggatccctg gtgcaacaca catcgccgca caagaccgtt tcaaacagg 540	
	tctgaggaag ctgggagggg ttctaggtcc cgggtctggg tggctgggga cactggggag 600	
55	gggtgtgctc tccctgggt ccttatgggt ggggtgggac ttggccgag tgcacttctc 660	
	gactgtctcc catgtgtcc ccgccag	687
	<210> 8	
	<211> 494	
60	<212> DNA	
	<213> Homo sapiens	
	<400> 8	
	gtgggtgcgc gggaccccc tgagacagccc tgctggacct tgggagtggc tgccctgatt 60	
65	gcacctcatg ttgggtggag gaggtaetcc tggttggggc gcaggaggag caggtgacc 120	
	tgctcaactgt gaggaacacac ctggacacta ggttggagg ettcagactt tctgtgagc 180	
	catggggcg actgtgacc ctgactgccc gggctctat tcccaaggag ggttccactg 240	
	gattccagtt tccgtctagag aaggaaacgc accgctctag ccaccagcc cgggtgctct 300	
	gcaccccagc cctgagccag ggttctcctg tcttgaggtc cagagagggg acacagcccg 360	
70	cctctgcctt ggggtctgga gtgtgtgggg tcagagagag agtgggggac accccagag 420	
	caggccctga gggcagaggt gatgtctgag ttctgcgtg gccactgtca gttcctctgc 480	
	ctccactcac acag	494
	<210> 9	
	<211> 865	
75	<212> DNA	
	<213> Homo sapiens	

5	<400> 9	gtaaaggttca	cggtgtgatag	tctgtgtccag	gatgtgtgtct	tctgggatat	gaatgtgtct	60
		agaatgcagct	cggtgtctgtg	atgcgtttctt	gtgggtggagg	tacttccatg	attttacaat	120
		ctgtgatattg	cggtgtgtggc	acgtgtgtgtg	cggtgtgtcat	gtatctgtgc	cggtgcaatt	180
		tgtgtgtgtgt	gtgtgtgtgtg	caacgtgtgtg	tccatgggtgt	gtgtgacctgt	gggtgtgcatt	240
		tgtgtgtgtct	tgtgtgtgtct	gcattgttcat	gctgtgtgct	gcattgtgct	gatgtgccta	300
		tttgtgtgtg	gtgtgtgtcat	gtgtgtgtgtg	catatgtgtg	tctatagcat	gggtgtgtgt	360
		ggccctcttg	ccttactcct	tctctctcca	ggcatgtgtc	gaaccattgt	ctctaccgct	420
		tctgtgtgtg	ctgtgtgggg	ctccacattc	agggttccca	ctcttagcat	gggtgcacct	480
		gtccctgtcac	agggtctggg	cttggagact	gtgaagccag	tttgagagaa	gaagttaggt	540
		gtctgtgtgta	ccttctctga	ccctctggac	ccccaggacc	ccagctctgc	ctatgccggg	600
		tgctgtgagat	atagggaagg	tgattctagg	ctctgtcccc	gggaacacat	cctccagagat	660
		cggccggggg	ccttgggggt	cgacgggggt	cgaggtctlg	gttcccaacc	720	
		acgtgtcatg	agcagcctgg	agggtgaagc	cttcaaaagt	gtgccaggcc	gggtgtcgaga	780
		gtgtgaagaag	tatccctgga	gcttctgtct	gggggagagg	acattgtgaa	accaccaagg	840
		acctcttctt	ctgacttctt	gagct				865
20	<210> 10							
	<211> 3782							
	<212> DNA							
	<213> Homo sapiens							
25	<400> 10	tgtgggatca	gttttcagtg	gtgggatagg	tggggatctg	tgggatttgt	ttttatgagt	60
		gggttcaaac	agggttcaag	gggtctcttc	tctcttgagt	gggtcttgag	gtgtgtccac	120
		aggtttattg	aggagaccat	atcttctctt	gaactatgtg	cggtgttata	gtaagtcagg	180
		gggtgtggagg	cctccccctg	gtctccctgt	ctgtttcttc	caactctggg	tctgtgttgt	240
		ctctgtctgtg	tgtgtggcgg	gtggggcagg	cttccaggcc	tcttctgtgt	caattggcctg	300
		tgtgtgtgccc	tgtgtacgct	ccgtctcttg	aattccccct	cgagtgtgag	gtcttctttt	360
		tttctttttt	tttttttttt	tttttttt	tgataacaga	gtctctgctt	tttttgccta	420
		ggctgtgagt	gttctggctg	atcttggctc	actgtcaact	gtgtgtcttg	agttcaaaat	480
		attctcttgc	ctcagcctcc	caagtgtgct	gaattatagg	cgccccacac	catgtcgact	540
		aatttttcta	atttttagtag	agacagagtt	tcttcaagtt	ggccaggctg	tcttcgaagt	600
		ctcgactcca	ggtgatccct	ccactctggc	ctccccaaag	gctgggagat	caggatgtgaa	660
		ccgcgcgcgc	cggccgagaa	tcgcttctct	caagttctct	gagatctctga	cgagatagct	720
		ctcgcagcct	tgggtctgag	aactctctct	aggtctctct	aggggtctct	780	
		ccatttctat	actctcttca	cagaagaggt	tcacgtgtgt	gtatttctcc	gtgtgttctt	840
		cggtaatattg	tgtctgtctg	ttatcgatgg	ctctcttcca	ttctctttag	ctgttgttta	900
	ttgttgtttt	tccggtctct	tgaagggaaa	gtttcgatta	tggatgtttt	aacttctttt	960	
	tcaaaaacag	catctgaagt	tgccgttttc	ctcttcaagg	agggatctcc	agggccccct	1020	
	gtgttgagtg	gcacccgtct	ggggccttgt	aggaaccctg	cgacacaggg	gaggttaggt	1080	
	gggtgtgtgg	gagccagctg	tccgcctctg	gcctccgcct	tctcagatca	cgagtggaat	1140	
	gcgtgtgtctca	gagcgcgaca	caacctactg	agaaactgtg	gtgagagggt	tctagattct	1200	
	gtgtcttcta	tgggaattca	atgcctgatg	atctgagggt	gaaccgtttt	ctccccaaag	1260	
	ctcccccttc	cccactgtct	tctgttgaaa	aaatctgtct	ccacgaaaac	agttccctag	1320	
	accacaagat	ttggggaccc	tgtgtcaaa	acctgtctca	cgactctctc	gtcagttgtt	1380	
	ataatattgct	ttttctgtgt	tgaagtcaga	ataaataggt	atttctgtga	tgtcttctgc	1440	
	cgactctcaga	cccactgggt	atttgtgggc	gtgtgtgctg	ctctctaggt	gggtgtgtgt	1500	
	caggccccat	gtaccttctt	gttactgtct	tccaggttgt	ttctcagggt	tgaattctga	1560	
	tcagatgtgtg	tttatcccaa	ggccctggcg	ccagctctct	ggggctgggg	aacatctgta	1620	
	agcacaagat	ccacgtgtgc	gtcttttgat	gccttcaaac	ctcagggcct	ctctgtgtctg	1680	
	tgttaagtgt	tgtctacgat	ctgtctcaat	ctctctcttg	ggaagccagg	gtcttagcga	1740	
	tccgttagta	aatgcagaga	actctggggg	agttctcgaa	ataggagagg	gggggtgggt	1800	
	ctctctctcc	cggctctctg	actcttctca	tgcctctgct	gtgtgtgca	ctgcatctcc	1860	
	gcaatctctc	cagcactggg	ctggagaggc	ccggagagct	gagtgccact	tgtgccacct	1920	
	gacgtgtgat	ggcagtcgtg	cacggggggt	tgatgtgtgt	tgactgtgga	tggcggtgtg	1980	
	tcaacaggggt	gtgatgtgtg	gtgactgtgt	atggcggtgt	tggggctctg	gtgggtgact	2040	
	gtggatggcg	gtcgtggggg	ctgatgtgtg	gtgactgtgt	atgacggctg	tggggctctga	2100	
	gttgtgtgact	gtgggtggcg	gtcgtggggg	ctgatgtgtg	gactgtgagt	ggcggtcgta	2160	
	gggtctgatg	gggtgactgt	ggatggcagat	cgtagggctg	gatgtgtgtg	gactgtgagt	2220	
	ggcgtgtctg	gggtctgatg	tgtgtgactg	ggatggcagat	cgtggggctc	gatgtgtgtg	2280	
	gacgtgtgat	ggcgtctgtg	gggtctgatg	tgtgtgtgag	ctggatggcg	ctgtctgggtg	2340	
	ctgatgtgtg	gtgatctgtg	atggcggtct	tggggctctg	tgtgtgtgta	ctgtgtgatg	2400	
	ggcgtgtgtg	ctgatctgtg	gtgatctgtg	atggcggtct	tggggctctg	tgtgtgtgta	2460	
	ctgtgtgtg	gtgtgtgtg	caggggtctg	atgtgtgtg	actgtgtgag	ctgtgtgtgt	2520	
	ggctctgatg	gtgtgtgact	tgaatgtgtg	tgggcaaac	gggtctgatg	tgtgtgtgact	2580	
	gtgtgtgtgt	tgtgtgggtg	ctgatctgtg	gtgactgtgt	atggcggtct	tgtccggggg	2640	
	cttgatgtgt	ggtgactgtg	gatggcgatc	gggtcacagg	gtctgtgtgt	tggctgactg	2700	
	ggatggcggt	cgtgggtctc	gatgtgtgtg	gactgtgtgt	ggcgtgtctg	gtgtctgtgt	2760	
	tgtgtgtgact	gtggatggcg	gtcgtggggg	ctgatgtgtg	gactgtgtgt	ggcgtgtgtg	2820	
	gggtctgatg	cctggggggt	tgatgtgtgt	tgactgtgtg	tggcggtgtg	gggtctgtgt	2880	
	gtgtgtgact	tggatggcag	tctgtggggt	tgatgtgtgt	tgaactgtgtg	tggcggtgtg	2940	
	gggggtctgag	tgtgtgtgtg	tgtggtggcg	gggtctgtgt	tctgatgtgt	ggatgactgt	3000	
	gatggcggtc	gtgggggtct	atgtgtgtgt	actgtgtgtg	cggtgtgtgt	gggtctgtgt	3060	
	gtgactgtgt	gatggcggtc	gtgggggtct	atgtgtgtgt	actgtgtgtg	gtgactgtgt	3120	
	acaggggtct	gatgtgtgtg	gactgtgtgt	gactgtgtgt	gggtctgtgt	tgtgtgtgtg	3180	
	gtggatggcg	gtcgtggggg	ctgatgtgtgt	gactgtgtgt	ggcgtgtgtg	gggtctgtgt	3240	
							3300	

	gtgtgtgact	gtggatggcg	gtcgtagggt	ctgatgtgtg	gtgactgtgg	atggcagtcg	3360
	gtccagagag	tctgatgatg	ggtagacttg	gatgggggtc	gtgggtcttg	atgtgtgttg	3420
	actgtgagtg	gcggctcgtgg	ggctctgatgt	gtgggtgactg	tggatggcgg	tcgtggggtc	3480
	tgtatgtgtg	tgcattgtga	tggcggctgt	ggggtctgat	gtgagtactg	tggatgtgtg	3540
5	tcggtcacag	gggtctgatg	tgtggtagct	cgaggttgag	tcccagggtg	gtctgtagctg	3600
	acttttgcctc	ctgggcccc	cgggccccct	tcccacaaac	gaagctctcc	aggcgtcttc	3660
	tgggtctcat	ccgcgcacat	ggctctggcg	caggtccaca	cgtcctgate	ggaaagaaaa	3720
	agtgcccagc	cttggccggg	gcaggccaca	tttgtggctc	atgccctctc	ctctgcggcg	3780
	ag						3782
10	<210>	11					
	<211>	980					
	<212>	DNA					
	<213>	Homo sapiens					
15	<400>	11					
	gtctggggcac	tgccctgcag	ggttggggac	ggactccacg	cagtgaggtc	tccctgggce	60
	atcactctggg	ctcatgaccg	gacagactgt	tggccctggg	gggcagtggg	gggaatggac	120
	tgtgatgggg	gcatagtgag	ctgtgtgcct	tggcgaaatc	tgagctgggc	catgccagge	180
20	tgcgacagct	gttgcattca	ggcaccgtgt	caagcttctc	ggcgaggctc	ctctccagtt	240
	ccgcagtgcc	tttgtctcat	atttgtctaa	tgtctctctc	gccagttttg	actctgaggc	300
	caaaagaaag	gtgtccccc	cccttaggag	ggcaggccat	gtttgagccg	tgtcctgccc	360
	agctgggcccc	ttagtctgtg	gtctgaggcc	aaaggaacac	tgtccccctt	cttagaggga	420
	cggtcccgatg	ttgagccacg	cccccgctac	ggggcctctc	agtgctgggt	tgttccacgt	480
25	ggccctctgtg	ccctttgcag	atgtgggtctg	tcccagtggt	ctgtggagtc	tttgcagattg	540
	ccagtgagca	atctgcggag	tttaggggtg	agctcagtag	gtccagtgag	gtccagtgag	600
	ctctggggca	atttctcttg	ctcccagggt	gggggtggag	gtggctctgg	ctcgtgggag	660
	ccagaccctg	tgtccggcag	ctggggcagca	actcctggat	cacatatgac	atcccgggca	720
	cggtgggctg	tgtgggtgtg	agccccagctg	gacccacagg	tggcccagag	gagacgttct	780
30	gtgtcacaca	ctctgectaa	gcccatgtgt	gtctgcagag	actcggcccg	gccagcccac	840
	gatggccctg	cattccagcc	gagccccgca	cttcacaca	aaacactgag	ccaaaaggga	900
	cggaaggtct	tggccacgtg	gtcctgctgt	tctcagacc	caacgggtca	ctcccatgta	960
	tctcccgctc	gcttttgcag					980
35	<210>	12					
	<211>	2485					
	<212>	DNA					
	<213>	Homo sapiens					
40	<400>	12					
	gtgagtccagg	tggccagggt	ccattgcctt	ggcgggtgct	ggcggggctg	gcagggtctc	60
	tgtctacctc	tctcctgccc	cttcccacat	gcncttcttc	ccggggccac	cagagtcttc	120
	ttttctgggc	ccccggccta	ccgggtctctg	gggtgcacgg	tccggagacc	ccggaacac	180
	ggctcggctt	ggcgagcccg	gagcgagagca	gggtgcacac	gaggcctgga	aatggcaagc	240
45	gggggttgga	gttgctctctg	ctgtgaggac	gagggggcgg	gggtgtgtct	gggtcagggt	300
	tgcgcgcgag	gtttagacct	gcagcttgtc	agctccaaat	tactactgac	cggtggacac	360
	cggtctccac	acgcttgtat	ctctctctcc	cgatacaaaa	ggattttatc	cgtattctcat	420
	tcctgtacct	gtctgttgag	ccccggagg	ggcggggctc	tctctctctg	gactagattt	480
50	ccactctgga	atctctggca	tttaggggtg	gtctctctct	gtccctggag	gtccctggag	540
	ccatggggca	ggcgccctgg	gagagctgcc	gtcacacagc	caactgggtga	gccacactca	600
	cggtgttgtag	ggccagatgc	ctgggtgccac	atcacgtctc	ctggatttta	agtaaaacca	660
	cacacctccc	gcaggccatc	tgcctgcgac	cttgtgtgtg	ctctggggaga	gtgttagcac	720
	cggtggaaat	cgtgcacact	caaggctcatc	agcaaggcta	tcgcgagta	ggtggaaagt	780
55	ggaggctctc	ctctgggattc	gtctctccag	gataaaggac	gtlgcacagc	tccggaaagt	840
	tttatttaaa	aatataacta	tttaattatg	caatataagt	aatcactaat	ggatctcgca	900
	attataaat	tttataaatg	ataattagaa	attataagta	gtacacacgt	tctggaaaaa	960
	caacaattgc	acatggcagc	agagtgaaat	ttggccgagg	gacacgtgtg	caactgtgtg	1020
	tgacggcgcc	ccaggcccac	agaattcgtc	gacaaaagct	actccccaga	gaagccacca	1080
	ccaggctctc	tgtgtgttgt	gaattttat	aaagtggatc	aagtaacagta	cgctccacgt	1140
60	gtggcaggcg	tttggggaat	gtgaggtgat	gactgggtcc	tcatgcccct	acagagacag	1200
	gtgtactgtg	ctgtctctct	cgagcaggcc	ggagagctct	ggagagctct	gtcccactat	1260
	tgtgtccagt	tggcctctga	ataaaaaagt	cttcaaaaac	gtttgcccc	aaaaactaga	1320
	acagagagag	tttcccatcc	catgtgtctc	caggggcgta	ctgtctctgc	tgtactgtct	1380
	gggctgtggc	gactcctaga	gttgtgttgt	gtgtcttgtg	gcataaaagt	cagtctctct	1440
65	gccccatcact	gtgatattctg	caccagcaag	gaagacctct	ttttttttct	tttctttttt	1500
	gctatttgaga	cggaaagctca	ctgtgtcttg	ctggaggttg	agtgcaggtg	cggtatctca	1560
	actcactgga	ctctgcgctt	cccgctggct	agctctctct	gtctctcagc	ctcccgagca	1620
	gctgagatga	caggccaccca	ccccctgcgc	ctggctaat	tttgtatttt	tagttagagag	1680
	gggtttttgc	aggtctggcc	aggtctggct	cgaaactctt	actcagattg	atccacacca	1740
70	ctcggccctc	caaaagtcgt	ggattacagag	tgtgacacat	caagccccag	cggaagagct	1800
	ctttttaaag	tgaccacacta	tagcgtctcc	cgaanataac	aggctcttgt	tttgcagtag	1860
	gctgcagagc	atctctggca	acagagagtg	gctctgtgtg	gctctggaga	tggctgaggg	1920
	tcggtgggca	ggcatgctct	ctgggtgcac	ctttaggctt	caacgggctc	ttctgtcttc	1980
	actgttttgtc	tgaaaaacgca	ccatttggcat	cttctgtttg	agattttctg	cttctcgttg	2040
	gtcatctgca	actatggggc	aaagtttgtat	ccgttggcgc	gcagcggtca	catgtagggt	2100
75	ctcatgactt	tcaactgtga	caaatctctt	gaaaaaaaaa	aaaggagctc	ggtagagact	2160
	taactctggg	tcaagttgtct	ggctctgtga	ataaaataca	agatttagaa	aaccttaagt	2220
	gaagaaaaac	tgtatgatct	agagcaaggga	tgtggctcac	cctgtggctg	gattctgttc	2280

	agccgcceca	gtgcgatggtg	agagtgggga	gcaggagattg	ttgtttcaga	ggcttcacat	2340
	ggatgtgttc	tsagtggttc	ggcggtctga	tggtagagct	gtcggtttgtg	tgatataggt	2400
	tctgtgtctg	ttgttggttc	gggtttgagt	tacgcatgtc	cagcacatgc	cctgcgcctg	2460
	tctcaactgt	gtcttccgcg	cccgag				2485
5	<210> 13						
	<211> 1984						
	<212> DNA						
	<213> Homo sapiens						
10	<400> 13						
	gtgagggcctc	ctcttcccca	gggggggcttg	gggtgggggtt	gatttgcctt	tgatgcattc	60
	agtgttaata	ttctctgtgt	ttctggagacc	atgactgtctc	ttgtcttgagg	aaccagacaa	120
15	gggttcagccg	cctctcttgg	atgaagccgg	acggagcggg	ttgcacagcc	tgaggaactgc	180
	gggtctccag	caggctctgt	ccagcgcgcca	tgctccagagg	cctcagggct	cagcagcgcg	240
	gagggcgcgt	gcctctcatg	atgagcatgt	gaattcaaca	ccgaggaagc	accacagctt	300
	ctgtcacgtc	accagagttc	cgttagggtc	cttggggaga	tggggcttgt	gcagcctgag	360
	gcccacatc	ttccagcagg	ccctcgacag	gtggcctgga	ctgggcgcct	cttcagccca	420
20	tgcccattcc	cacttgcagt	gggtctcacac	ccaaggagcc	acacacctaa	atatcgtgcc	480
	aaactaaagt	ggttcaactc	agctgtctt	tattgcagac	agttaccttt	tttttttaa	540
	tactttaagt	tctagggtac	atgtgcacga	cgtgcaggtt	agttacatat	gtatacatgt	600
	gccatgttgt	ttgtctgcac	ccattaaact	atcatttaca	ttaggtatat	ctctcaatgc	660
	tatccctccc	cactcccccc	atcccatgac	agggcctgggt	gtgtgatgtt	ccccacctgt	720
25	ttgtccaaagt	ttctcatgtt	tcagttccca	ctgtgtgagt	agaacatgtg	gtgttttgtt	780
	ttctttctct	gcaatagttt	gctccagagt	atggtttcca	ctctctgtca	tgctccctaca	840
	aaggaacatga	actcaactc	ttctatgagt	gcaatgtga	cctcaggtgt	tatgtgcgtg	900
	attttcttaa	ttccagttcat	catcgatgga	catttgggtt	gggtgcaggt	ctttgtctact	960
30	atgaataagt	ccgcaataaa	catacgtgtg	catgtgtctt	tatagacaga	ttgtttatac	1020
	tctttggagt	atataccacg	taattgggag	gctgggttga	atggttattc	tagttctaga	1080
	tctcttgagga	atccaccac	ttcttccac	aatgggttga	ctagtttaca	ctccaccaca	1140
	cagtgtaaaa	ttgttcttgt	gctggagagg	atgtggagac	cagttatttt	tttatgaata	1200
	tagtatcatc	gaacaaagc	acagttagtg	aaggatgggt	caggaaagct	gcaggccaca	1260
35	cagcatcttc	ttccgaagac	tccgggtttt	ttctgtgcat	cttttgaac	tttagctcca	1320
	attatagcat	gtacagtgga	tcaaggttct	ttctcatata	ggttcaagtt	ctagatgtaa	1380
	ataagtttgt	gtaacagaaa	caaaaatttc	ttgtacacac	aacttgctct	gggatgtgga	1440
	ggaaagtgct	ctcgagctgg	cggcacacgt	gtcacccctc	tgggagacaga	tacctctgpc	1500
	ccatggatcat	ggggcgctgg	gcttggggag	gtgggttaca	cagtcgacac	tgaccgtggg	1560
40	cctgtggata	tttctgagct	ctcgagatct	ctcgagagac	acagctgcga	ttgtgttgata	1620
	ggggccacag	ttgttcagag	ggggcgaggt	tccacgcccc	agctttctta	ccgtcttcac	1680
	ttatttttcc	tcaagagctc	gagaagttgg	gcggcgctcg	atggcctctg	ttcgtctcca	1740
	gctgggcacag	aattgcacaa	gctgatggta	aacactgagt	acctataagt	aatgaggaat	1800
	ttcgtatagca	gttaactgtg	gagagctcgt	ctgtttggaaa	gaataattag	ttttctattt	1860
	aaccgctttg	gagaatgtta	ctttatttat	ggctgtgttaa	attgtgtgac	attcagttcc	1920
45	ttcgtagcacg	atactacgta	aaaaagttaa	agttaacctt	gctgtgtatt	ttcccttatt	1980
	ttag						1984
	<210> 14						
	<211> 1871						
	<212> DNA						
	<213> Homo sapiens						
50	<400> 14						
	gtgagggcccg	ttccgtgtgt	ctgtggggag	ctccacagcc	ttgtgggcttt	gcagttgagc	60
	ccccctgtct	ctgcctctgg	acccgcagcg	ttgtctctgc	caggtccctc	ctctctcgccg	120
55	gtgtcggatc	cgcaagagtc	gaggcctctg	gctgtgtgac	caggcctagg	ggcgacgggg	180
	ccactctggg	aggaggtggg	taccgtgcag	gcctctgttc	tgacagagag	caccacaggt	240
	acacacagag	ttagtgccag	cgttgacctg	gtcctctgct	ctctttggaa	gtacaagagt	300
	ggcgctctct	ggggccccag	ttgagacccc	aggagctgtg	cacagggcct	gcaggggcga	360
	ggcggcagcc	ttctccccag	gggtgcacct	agcctcgga	gagcagagag	ttgcttagtga	420
60	gtgtggccac	agcgttcgt	gcgttcacct	ttgtctctgc	gggtgtggag	gtgtgtgtgc	480
	agatattgga	tttctgtgag	gctgtgtctc	tgacacacag	agatgtgttc	gagtggtttt	540
	cagagttgat	ttttgtgaat	caaaactaaa	tcaggcacag	gggaactctg	ctcagcacag	600
	gggatgttcc	aattgtgttc	ccctcaaggg	gcgccacag	agccggtggg	ctgtttttaa	660
	agtcgcatct	gacgagggag	gagaaaacct	gaagactgtg	aggggaacct	tcagaaaatt	720
65	tgagcccccag	gggtggtttc	aggtgctttg	ctggggtctg	tttgtgaaaa	cccatcttga	780
	cccgctctcc	agtcgacac	tcagagttca	ccctccaggg	ccgctggggt	ctgggggtat	840
	gctctgggtt	ctttgtgccg	cagccgggag	caagacagg	ttgtgcacct	taaaacctac	900
	aagattctac	ggggggggag	ccaggtccca	agcaactgag	ggctcaggag	ttctgagggt	960
	gctgagggga	cagagcagac	ggggaacgct	gcttctgtgt	gcgaagttcc	ttgaggtgct	1020
70	ggccacgggag	gtgcttcaga	gtgtatgttg	gggtccccc	gggggcagaa	ctctgtcttc	1080
	gtatgactcg	cagccatgta	acaggaaggg	gtggccacag	ggaagctgga	atgcaccagg	1140
	ggagctcgag	agctggcgag	gggtcccgcc	aggaaggaaa	gggatgcc	agggccagag	1200
	ggggccacag	ggagctcccg	agctgggtga	gcgaggtctg	tgactcgccg	agcgaaacct	1260
	gggcacaggg	gggtctccct	tcgtgtgtgc	cagctcacag	ccccagcagg	ttcccgccct	1320
75	cttgacgtga	agctgacgac	ccctttgtc	taaaagacag	cagatgcctt	ggcgacatct	1380
	gagcagaaac	agtcacacct	gttgagaaac	gctctaaaac	agatgggagt	gggtgggaat	1440
	aggagaaaaa	agggaaagtc	gttgagaaac	gctctaaaac	agatgggagt	gggtgggaat	1500
	ttctgtccag	atttttagtt	gccccgagac	acagataggt	ctataacagg	actgtgggtt	1560

	actgcacac	acaagcac	acacagacat	gcattgatgc	atccgtgtgt	gtgcacctgt	3060
	gcccattgag	aaacccatgc	atgtgcatc	acagggcacc	gtgggcccat	ggagggccat	3120
	gcccacaccc	acgagcaccg	tctgattagg	aggcctttcc	tctgacgtg	tcgcgcattc	3180
	tctcag						3186
5	<210> 18						
	<211> 781						
	<212> DNA						
	<213> Homo sapiens						
10	<400> 18						
	gtatgtgcag	gtgcctggcc	tcagtggcag	cagtgcctgc	ctctgtgtgt	tagtgtgtca	60
	ggagactgag	tgaaatctggg	cttaggaagt	tcttaccctt	tttgcgtaca	ggaaagtgtt	120
	taaccacacc	actgtcaggc	tcgtctgccc	gccctctcgt	gggttgagca	ggacacctga	180
15	tggaaggagc	aggagctgtc	tggagctgc	cactctctcc	acctgtctct	gctctgggaa	240
	gcgtctgggg	gcctgggtctc	tctgttttgc	cccatggtgg	gatttggggg	gctctggctc	300
	tccatgtttg	cctgtggtgg	gattgggctg	tctccctctc	atggacacta	gggcccctgt	360
	gcacaaccag	gcgaagggtc	taggagagg	cagagccacg	gctacccacg	cctctctcag	420
	agcagagccc	gcgtatcaac	acgacagagc	cccgtgcttc	ccagtcaacc	ccagtcaacc	480
20	tctctgtccc	ctggacacct	tgtccagcat	caggagggtt	tctgactcgt	ctgaattaca	540
	agccatgtcg	aacctggcgt	cctgagctta	acagcttcta	tcttctgttc	ttctgtgttc	600
	gtgaaatttt	cacctggaga	agccgaagaa	aacattttct	tcgtgactcc	tgcgggtgctt	660
	gggtctggag	agccagagat	ggagccaccc	cgcagacgtc	cgaggtgggg	cagctttctg	720
25	gtgtctctctg	ggagggggagc	tgggctgggc	ctgtgactcc	tccgctctctg	ttttccccc	780
	g						781
	<210> 19						
	<211> 536						
	<212> DNA						
	<213> Homo sapiens						
30	<400> 19						
	gcaagtgtgg	gtggaggcca	gtgcggggccc	cacctgccca	gggggtcatcc	ttgaacagccc	60
	tgtgtggggc	gagcagccctc	agatgctgct	gaagtacaga	cgcccccggg	cctgacctgt	120
	gggggtcaga	gccacgcggg	agctctggag	tgatataagc	ctgggtctcc	cagggtctcc	180
	agcctgtggc	gttccccaac	ttcttgaaac	cctgcttccc	atccaggggg	cgatggcttc	240
	ccaagctgtg	gagcctttctg	acccttgacc	tgtgtctctc	ccacagctct	tctctggtct	300
	ctgcctctag	ctctctgggtt	cctgagcaag	ttctctctcc	ccggcccgcc	tccaagctca	360
40	ctgggctgcc	tgttctgctg	ccccggtaga	ggggtgtctg	tcccttaacc	gaggttccca	420
	ctcagcaggg	ccacagagtg	caggccctgc	ctgcctcgcc	accacacagt	cctagagggg	480
	ttggaggatg	ccactctggg	cctctttctg	aacggagctc	gattttggcc	ccagag	536
	<210> 20						
	<211> 3179						
	<212> DNA						
	<213> Homo sapiens						
45	<400> 20						
	atctcatgtt	tgaaatctaa	tgtgcactgc	atagacacca	ctgtatgcga	ttacagaagc	60
	ctgtgagatga	acgggggtggt	ggtcagtgcc	ggccccatggc	ctggctgtgc	atttacggaa	120
	gtctatgagt	gaatgggggtt	gtgttcagtg	cgggcccatg	gccttgctgg	gacctggagg	180
	ttctctatgc	tgtgagggag	gaggggaagg	aggttagggg	atagacagtg	ggagcccca	240
	ccctggagaa	caataacagta	agtcacagcc	cgaagggtgc	cagggatgct	ggggcccgac	300
55	cttggggcgg	ggggatgatg	gagggcctgg	cagggtgtgc	agggatgagc	ggggccccac	360
	ctgggggtggc	agggtgtgatg	gggggggtgc	gtctgggttg	cggggaagat	ggggaaagctt	420
	gggtctggccc	ctctctcccc	tgctctccac	ctgcagcgtg	ggatccggat	gtgcttcccc	480
	gggtcacacc	ctctggggcca	tcagctttca	tggagggtgg	gggcaggggg	atgacacatt	540
	ctctgtataa	atccagagtt	ctctctcttg	aaagcccaaa	ctcaggttga	aagtcacatt	600
	cgccctctgg	ccattctctt	aagagtagac	caggattctg	atctctgaag	ggtgggttagg	660
60	gtggggcagt	ggagggtgtg	gacacaggag	gcttcagggt	ggggctgtgt	atgctctctc	720
	atctctttat	catctcccag	tctcatctct	catctcttta	tcatctccca	gtctcatctg	780
	tctctctctt	atctctccagt	ctcatctgtc	atctcttcca	catctcccag	tctctctctc	840
	tatctctctt	tctctctagtc	tctctccagc	tactctccca	ggggcggtgc	caggctcgca	900
65	gtggagctgg	acataagcttc	ttctctcagg	agaaggaact	ggaagggttg	cagagaacag	960
	gaaggggcgg	tcagagggaac	gcagttcttg	ggtgaagaaa	cagccccctc	tcaagaattg	1020
	gcttggggcca	ccagaaaaccg	agggccctgc	gtgagttggc	ccagagacct	ccagcagctc	1080
	cctgtgtggg	ccttatggta	tggccgggtc	ctactgagtg	caacctggac	agggtctctg	1140
	gcttgagctg	agcccgagcg	tgcctgggtg	cggggtgggg	gcttatggcc	actggatagc	1200
70	ggctcatctta	ttgctgctgc	ttcagagaa	gtctgagtg	ccagacttaa	tgtgagatgt	1260
	gggcccaggt	ccacagcctg	tgtcgtataa	gcactctgtg	ctctggagcc	ccgttatagg	1320
	agctgtgagg	aaaggagggc	tcttggcagc	eggctctggg	ggccttttgc	cctgcaaaact	1380
	ggaaggagagc	ggccccggggc	ggcttggggc	gagcagctca	agtgaagagt	tggacagaca	1440
	aggggggggga	tctcccaggga	gcagaggcgc	ctgctcaggc	acacctgggt	ttgaatacaa	1500
75	gaccacaagg	tcaggccatt	gttcagctat	ccatctctta	caaaagctca	gattctctgt	1560
	ttcccggttg	ttttctgttg	aaattttact	caggtattct	tattattttt	gtctaaagat	1620
	cagacactta	aaaaaggtat	tgtctttgat	atgctttaac	tactaagctc	actactttat	1680
	ttgtctcttt	ttatttatta	ttattattat	tattagagat	ggtgtctact	ctgtcaccca	1740

	gggtgttagt	gcagtgccac	agtcattggt	cgtgtagcc	gcaaaccccc	aggtctcaagt	1800
	gattctccgg	cttcagcttc	ccagagtgc	gggattacag	gtgtgagcca	ctgcccttgc	1860
	ctggaccctt	tataaacccac	tatgtaaggt	caggtccagt	ggcttcacaca	cctgtcatcc	1920
5	cagttagttt	ggaagccgag	gcagaaggat	tgcttgaggc	caggagtttg	agaccagcat	1980
	gggtaacata	gggagcccc	atctctacaa	aaaatgcaaa	aagttatccg	ggcgtgggg	2040
	ccagcatctg	tagtcccagc	tgctcgggag	gctgagtggg	aggatcgctt	gagcccgga	2100
	ggtcatggct	gcagtgaagc	gtgattgtac	catcgactc	cagcctgggc	aacagagtga	2160
	gacctgtgt	caaaaaaaa	aaaaaanaa	gaaggagag	gagaaagaa	gaagaaggaa	2220
	gaaggaaa	gaagaaga	gaagaaggaa	gaagaaggaa	gaaggaggcc	tgctaggtgc	2280
10	taggtcagct	gtcaaatctc	agagcaaaat	gaaataaaca	aagttttaa	gggaaagaaa	2340
	aacccagct	ctttggactt	ccttaggcct	gaacttcac	tcaagcagct	tccttcaca	2400
	gacaagcgtg	tatggagcga	gtgagttcaa	agcagaaagg	gaggagaagg	aggcaagggt	2460
	ggaggctgtg	ggtgacacca	gccaggaccc	ctgaaaggga	gtgggtgttt	tcctgcttca	2520
15	gccccagct	ctgcccgtc	ctgcaactgc	tgtaaccgtc	gatgttggtg	ccaggtgcc	2580
	acctgggaag	gatgctgtgc	agggggcttg	ccaaactttg	tggggtttca	gaagcccgag	2640
	gcacttgtgg	caggcaccaat	tacagccctc	ccccaaagt	gccccagctc	ttctcctgga	2700
	acctgtgaat	gtgtcacccg	caaggcagag	gctgggtgaag	gctgcaggtg	gaatcacggc	2760
	tgccagtcag	ccgatcttaa	ggtcatcctg	gattatctgg	tgggcctgat	atggccacaa	2820
20	gggtccctag	aagtgaagaga	gggaggcagg	ggagagtcag	agaggggacg	tgagaaggac	2880
	cactggccac	tgtgtgctt	gagatggagg	aggggttccc	cagccaagga	atggggcgag	2940
	ccgtctccatg	ctggaanaag	aagcaatcct	ccccgtctct	gagggcacac	ggcctgcc	3000
	acgcctcgat	ttcaggccag	tgggacctgt	ttcagctt	cggcctccag	agctgtaaga	3060
	tgatgcgttt	gtgttcagcc	actaagctgc	agtgattcgt	cacagcagca	aatggaatag	3120
25	cagtacaggg	aaatgaatc	agggacagtt	ctcagagtga	ctctcagccc	accctctggg	3179

THE UNIVERSITY OF CHICAGO

- THE UNIVERSITY OF CHICAGO

9. Process for identifying substances which affect the promoter activity, silencer activity or enhancer activity of the human catalytic telomerase subunit, comprising the following steps:
- 5 A. adding a candidate substance to a host cell which harbours DNA sequences according to one of Claims 1 to 3, which sequences are functionally linked to a reporter gene, and
- B. measuring the effect of the substance on expression of the reporter gene.
- 10 10. Process for identifying factors which bind specifically to the DNA according to one of Claims 1 to 3, or to fragments thereof, characterized in that an expression cDNA library is screened using a DNA sequence according to one
- 15 of Claims 1 to 3, or subfragments of widely differing length, as the probe.
11. Transgenic animals which harbour recombinant constructs or vectors according to Claims 4 to 6.
- 20 12. Process for detecting telomerase-associated conditions in a patient, comprising the following steps:
- A. incubating a recombinant construct or vector according to Claims 4 to 6, which additionally contains a reporter gene, with body fluids or cell
- 25 samples,
- B. detecting the activity of the reporter gene in order to obtain a diagnostic value, and

- C. comparing the diagnostic value with standard values for the reporter gene construct in standardized normal cells or body fluids of the same type as the test sample.

Regulatory DNA sequences of the gene for the human catalytic telomerase subunit, and their diagnostic and therapeutic use

A b s t r a c t

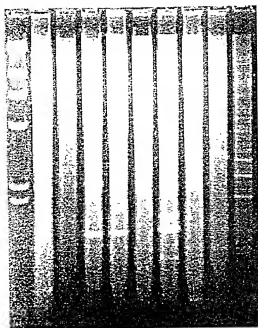
This invention relates to regulatory DNA sequences, comprising promoter sequences and intron sequences, for the gene for the human catalytic telomerase subunit. In addition, this invention relates to the use of these DNA sequences for pharmaceutical, diagnostic and therapeutic purposes, especially in the treatment of cancer and ageing.

001260.9428464

Fig. 1

A

1 2 3 4 5 6 7 8 9 10



B

1 2 3 4 5 6 7 8 9 10



Fig. 2

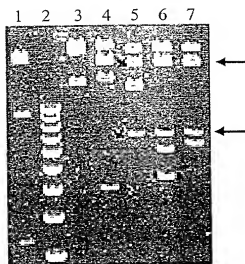
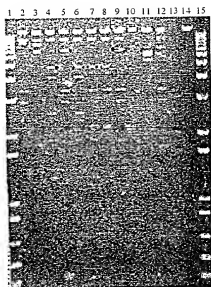


Fig. 3

A



B

2 3 4 5 6 7 8 9 10 11 12 13 14



Fig. 4

GAGCTCTGAA	CGCTGGAAAC	GAACATGACC	CTTGCTCGCC	TGCTTCCCTG	GGTGGGTCAA	GGTAATGAA	70
TGCTGTGCGA	GGAAATGGCC	ATGTAAATTA	CACGACTCTG	CTGATGGGGG	CAAGTTCCTC	CATCATATT	140
CATCTTCCGC	CCCAGGACTG	GAATGATGCC	AGCAACTCTG	TCGGGTGTGA	CAAGCTATGA	CAAACTCATG	210
TACAAACACC	ACTCTTTTAC	TAGGCCCAAC	GAGCAGGGGC	CACACCCCTG	ATATATTAG	AGTCCAGGAG	280
AGATGAGGCT	CGTTTTCAGC	AGCAGGCTGG	GGTGACAACA	CGGGCTGAAC	AGTCTGTCTC	TCGTAGATCT	350
TAGACCTCTG	CAGCGATCTC	CCCAAAATCT	AGGGCCCTGT	TGCTGCTTCC	CGAGGGCCGC	ATCTGCCCTC	420
GGAGCTCAGC	CTGGGGTGGC	ACACTGAGGC	CAGCCCTCTG	TCCACACCTC	CGCCCTCGCC	GCCTCAGCTT	490
CTCCAGCAGC	TTCTTAAACG	CTGCTTGGGC	CGTGCTCCAG	CGCTACTGTC	TCACCTGTCC	CACGTGTGCT	560
TGCTCTAGCG	ACGTGATCTC	CAGGCTTCTC	CTCCATATGG	GGTGTCTGTC	TCTCTCCACA	ACACTGTACT	630
CGCTTGAAGG	GAGGAGTCTC	TGCGGCTCCC	AGACTGGGCT	CTCTGAGGCT	GAACCTGGCT	CGTGGCCCCC	700
GATGCAAGTT	CTTGGCTCTC	GGCTGCACGC	TGACCTTCAT	TCCAGGCGGC	TCCCGCTCTC	TTGCATCTGT	770
CGCGGGCCCT	CGGCTGTGTT	CTTCTGTTTC	TGTGCTCTTT	TCCAGCTCCA	GCTGGCTGTG	TCCTGCCCCC	840
CTAGGCTCTC	CGGCTGTGTT	TAGGCTATGG	ACGGGGGCGT	GGTGGGCCAG	GGCGCTCTTG	GGAATTCGAA	910
CATTGTTGGT	TGAAGATTAG	AGTGCCTGTC	CTCAGCTAGG	TCCACGGGCA	CAGGCGTGGG	GATGGAGCCC	980
CCCGCAGGGA	CCGCGCCCTC	TCTGCCCAGC	ACTTTCTGCG	CCGCTCCCTC	TGCGAATCAA	GATGGCAGT	1050
TTCCACAAGC	ACTAAGCATC	CTCTCCCAA	AAGACCACGC	ATTGGCAGCC	TGCGACATT	GCCCCACAGC	1120
CCCTGGGAAT	CAGTGACTA	CGCACATCAT	GTACACATCT	CCGCTCCAGA	CCGACCCCTG	CTGTTTTATT	1190
TTAATGACTA	CAAGACAGGG	AAATCCCTGC	TAAATGTGTC	TTTAACAAC	TGGTTAACA	AACGGGTCAA	1260
TCGCGAGGCT	GGACACTTCC	TACAGTGAA	GAGGACAATG	CCGTTTATAA	AGCTGACAG	CACTCTAAGG	1330
GANTTAGCGT	GAGTCAAAAC	TGCCACTCTC	ATGGGATAGG	TACGCAACAT	GCTCAAAAG	GAAGAATTT	1400
ACCCCTACGCT	AGGGGAGTGG	TAGGACCGGT	TAGGACCGGT	GGGGGGGGCA	GCTGGGGGCT	ACTGACGCA	1470
CCCTTTACTA	AAGCAGTATT	CTCGGTTCTG	ATGGTATTGG	CTCGACTATG	GGAGACTAAC	CATAGGGGAG	1540
TGGGATGGGG	GGAAACCCGA	GGCTGTGCCA	TCTTTGCCAT	CCCGGAGTGT	CTCTGGACAG	ATAATGCTCT	1610
AGGATGCCCC	ACTGCTCATG	TCCGCCAAAC	CTGTGGACAG	AACCCGCGCG	GCCCCAGGGC	CTTTCAGGCT	1680
GTGATCTGGA	GAGGACACCT	AGGATCTGGG	ATCTTCTGGG	ACTACTCCAG	AGCCGGAAGA	GAATACGAG	1750
GGTCTGTGGG	AGAGCGCGGC	ACGGAGGTCA	GAGGGGGGCA	GCTCAGGAG	GATGGAGGCA	GTACGTCTGA	1820
GCTTGAAGAG	AGGCGGAGGG	CTCGAGCCC	AGGCTCGCAA	AGGCTCCGAG	GCGCTGTGAA	AAGCCGAGAA	1890
AGCCGCTCTC	ACGGAGCCTC	CAGCAGAAAG	GACAGGGCTG	CCCTTAGCCG	ACCAGGGCCC	ATCGTGGACC	1960
TCGCGCTCTC	GGCTCATAGG	AGGCACTGAG	CGCTGCCCTT	CTAGCATGAA	GTGTGTGGGG	ATTGTCAGAA	2030
GCACACAGAA	ACCCATGCAC	TGTGAATCTA	GGATTATTTC	AAAACAAGAG	TTTACAGAAA	CATCAAGGA	2100
CAGGCTGAAA	GCTGCTCCGG	GACAGGTCAG	GGCAGGTCAG	AGTGATTTTA	TTTAGCTATT	TATTTTATT	2170
TACTTACTTT	CTAGACACGA	GTTATGCTCT	TGTTGCCACG	GCTGGAGTGT	AGCGGCAATG	TCTTGCTGCA	2240
CTGCAACTCT	CGTCTCTGGG	TTCTCAGCAA	TTCTCGTGCC	TGACGCTCCC	AGCTGCTGGG	GATTTACGCG	2310
GTGCAACACC	ACACCCGGCT	AATTTGTAT	TTTTAGTAGA	GATGGGCTTT	CACCTCATTT	GTCAAGCTGA	2380
TCTCAAATCT	CGCTACTCAG	GGTATCCGGC	CACCTCAGCC	TCCCAAGAT	TCTGGATTAT	AGGCAATGAG	2450
CACTGCACCT	GGCTATTATA	ACCAATTTAA	AACTTCCCTG	GGCTCAAGTC	ACACCACTAG	GTAAAGGATT	2520
CATGGAAGTT	AAATTTCCCT	TTACTCAGGA	GTACCCCTCC	TTTGATATT	TCTGTAATTC	TGCTAGACTT	2590
GGGGATACAC	CGTCTCTTGA	CATATTACCA	GTTTCTGTGA	CCACTGTGTA	TCCCATGGGA	CCCATGACAG	2660
GGGCACGTGG	GAGGCTGTGA	GGTTCAGGTC	CAGTGGGGT	TGCCATCTGC	CAGTAGAAGC	CTGATGTAGA	2730
ATCAGGGCGC	AAGGTGTGAC	ACTGTGCTGA	ATCTCAATGT	CTCAGTGTGT	GCTGAAACTG	TGAGAAATTA	2800
AAGTCCATCC	CTGCTACTCT	ACTGGGATTG	AGCCCTCCCT	CTATCCGCCC	AGCGGGCAG	AGGAGTCTCT	2870
CTCACTCTCT	TGGAGGAGAG	ATGATACTTT	TGTTATTTT	CACCTGCTGT	ACTGAATCCA	CGTCTTCATT	2940
TGTTGGTTTG	TTTGTGTTG	TTTGAAGGCG	GGTTTCACT	TGTTGCTCA	GCTGGGAGGG	ATGCTAAATG	3010
CGGAGCTTTG	GCTTACTGCA	GCTCTGCTC	CCGAGGTTCA	AGTGATTCTC	CTGCTTCCGC	CTCCCAATTT	3080
GCTGGGATTA	TAGGCACCCG	CACCATGCCC	CAGCTAATTT	TTTGATTTT	TATAGAGAC	GGGGGTGGGT	3150
GGGGTTCAC	ATGTTTGGCA	GGTGGTCTC	GAACCTCTGA	CCTCAGATGA	TCCACTGCC	TCGCTCTCT	3220
AAAGTCTGG	GATTACAGT	GTGAGCCACC	ATGCCCAAG	CAGAATTTAC	TCTGTTTGA	ACACATCTGG	3290
TCGTAGGTAG	GAGGCTCACC	CACCTCAAGT	GTTGTGTGTT	TTTAAGCCAA	TGATGAATAT	TTTTTATTGT	3360
TGTTTAAACA	CTCTGATGCT	TTTACACTGT	GATGACTAAG	ACATCACTAG	CTTTTCAAG	ACACACTAAC	3430
CCATGCCATA	ATACTGGGGT	GGTCTCTGGG	TATCAGCAAT	CTTCATTGAA	TGCCGGGAGG	CGTTTCTCCG	3500
CCATGCACAT	GGGTGTTAAT	ACTCCAGCAT	AATCTTCTGC	TTCATTTCT	CTCTTCCCT	TTTTTAAAT	3570
TGCTTTTCT	ATGTTGGGCT	CTCTGAGAG	AACCAAGTGA	AGCTACAAGT	TAACTTTTGT	TGGAAATGAT	3640
TTTCCAAACC	CGCCCTTTGC	CCTAGTGCCA	GAGACAATCT	ACAAACAAC	CCCTTTAAA	AGGGCTTAGG	3710
ATCACTAAGG	GGATTCTTAG	AAGAGCGACC	TGTAATCTTA	AGTATTACA	AGACGAGGCT	AACTCCCGAG	3780
GAGCGGTGCA	CGCCAGGAGG	GGTGCAGGCG	CGTTCAAAAT	GCTAGCTCCA	TAAATAAAGC	ATTTTCTCTC	3850
GGCAGTCTCT	GAAATGAGGA	AAGGTTACAT	TTAAGGTTGC	GTTTGTTAGC	ATTTAGCTGT	TGCGGACCT	3920
CAGCTTACAT	ATCCCTCGGA	GAAGCTGGGA	GACCCAGAA	TTCTCGCCCC	CTTCTAGATC	AACTCTAGC	3990
AACCCGGAGT	CTGATTTCT	GGGAAGTCT	CAGCTGTCT	CGGTTGTGTC	CGGGGCCCA	GCTCTGAGAG	4060
GGACCATGGC	CGGTTGGGCT	TCTACTGCTG	GGCTGGAAGT	CGGCGCTCT	AGCTCTGAG	TCCGAGGCT	4130
GGAGCCAGGT	GCTGGAGCCG	CAGAGGCTGC	CTCCACCTGT	TGCGGGCGGG	ATGTGACGAC	ATGTTGGGCT	4200
CATCTGCCAC	ACAGAGTGCC	GGGGGCCAGG	GCTGAGGCGG	TTGTGGGTGT	TGTGAGCCGC	CGGTTGGGCG	4270
CCGAGCAGGA	CGGCTGGGCT	CGATTTCCTA	CCCTTTCTCG	ACGGGACCGC	CCGGTGGGCT	GATTAACAGCA	4340
TTTGGGTTGG	TTTGTCTCAT	TGCGGGACCC	CTCGCGGCT	GAAGAAGCTG	GAAGAAGCTG	CGCGGCGCTG	4410
TGTCAGGAGG	CCCAAGTCCG	GGGGGAAGTG	TGCAGGAGAG	CACCTCCGGA	GCTCGCGGCT	CGCGCTCCAG	4480
GGAGCAATGC	GTCTCTCGGT	TGCTCCCGAC	CGCGCTCTAC	CGCGCTCCGT	CTCTCCCTCT	ATCTCCGCGA	4550
TGCTGGTGGC	CGCGAGCCCG	ACGCCCGCGC	TGCGGACCTG	GAGCGAGCCG	TGGGCTCTCG	GATCAGGCA	4620
CGCGCACAA	GGTGCGCGCA	CGCACTGTT	CCAGGGGCTT	CACATCATG	GGCCCTCCCT	CGGTTTACCC	4690

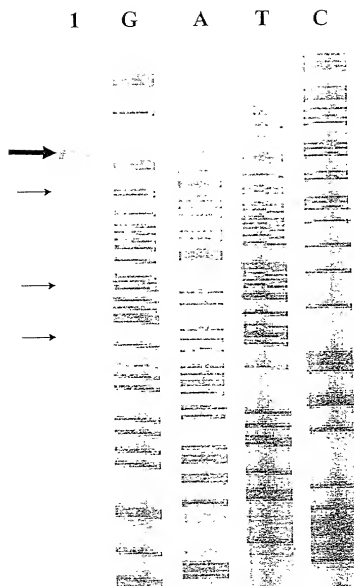
- 5/15 -

Fig.4 (continued)

CACAGCCTAG	GCCGATTCTGA	CCCTCTCTCG	CTGGGGCCCT	CGCTGGCGTC	CCTGCACCCT	GGGAGCGCGA	4760
GCGGCGCGCG	GGCGGGGAAG	CCCGCCCCAG	ACCCCCGGGT	CCGCCCGGAG	CAGCTGCGCT	GTGCGGGCCA	4830
GGCCGGGCTC	CCAGTGGATT	CCCGGGCACA	GACGCCCAGG	ACCGCGCTCC	CCACGTGGCG	GAGGGAATGG	4900
GGACCCGGGC	ACCCTCTCTG	CCCTTTCACC	TTCGAGCTCC	GCCTCCTCCG	CGCGGACCCC	GCCCGCTCCC	4970
GACCCCTCCC	GGTCCCCCGG	CCGAGCCCCC	TCCGAGCCCC	CCCTTCTCTT	TCCCGGGCCC		5040
CGCCCTCTCC	TGCGGGCGCG	AGTTTCAGGC	AGCGCTGCCT	CCTGCTGCGC	ACGTGGGAAG	CCCTGGCCCC	5110
GGCCACCCCC	<u>GCGATG</u>						5126

09/562246.092100

Fig. 5



- 7/15 -

Fig. 6

GTTTCAGGCA GCGCTGCGTC CTGCTGCGCA CGTGGGAAGC CCTGGCCCCG GCCACCCCGG CGATGCGCGG 70
 CGCTCCCCCG TGCGGAGCGG TGCGCTCCCT GCTGCGCAGC CACTACCGCG AGGTGCTGCG CTGGGCGCAG 140
 TTGCTGCGCG GCGCTGGGCG CAGAGGCTGG CGGCTGGTGC AGCGCGGGGA CCGCGCGGCT TTCCGCGCGC 210
 TGGTGGGCCA GTGCTGGTGG TGCTGCGCTT GGAAGCGACG CGCGCCCCCG CGCGCCCCCT CTTCTCGGCA 280
 GGTGTCTTGC CTGAAGAGCT TGCTGCGCCG AGTGTGTCAG AGGCTGTGCG AGCGCGGGCG GAAGAAGCTG 350
 CTGCGCTTGC GCTTTCGGCT CTGAGCAGCG CACTGCGGGG GAGCGGGGCG TGCGGCGTGC TGGTGGCGCG 420
 CGAGCTACCT GCGCAACAGG GTACACTGCT GGCACTGTC CGGCTTTTGG GTCTGGTGGG TCCTGCTGCG 560
 CGTGGGCGCG GACGCTGTGG TGCGGCGGCG CCGCTGTAC CAGCTCGGCG CTGCGACTCA GGCCCGGGCG CGCGCAGCG 630
 GCGTACGAGG TGTGCGGGG GCGCTGATCG AACCGGCGTG GAACCATAGG GTACAGGAGG CGCGGCTGCC 700
 CTAGTGGAGC CCGAAGGCGT GTGCGAGGAG GCGCGGGGCG AGTGCCAGCG GAGTCTGCGG GTTGCCCAAG 770
 GCGTACGAGG TGTGCGGGG CCGTGAAGCG GAGCGGACG CGGTTGGGCA GGGGTCTGCG GCCACCCCGG 840
 AGGCGCAGCG GTGGCGCTCG GCGCTGATCG GAGCGGACG CCGTGGGCG AGACCCCGCG AAGAAGCGAC 910
 GCAGGACGCG TGAAGCGAGT GACCTGTGTT TCTGTGTGGT GTCACTGCG AGACCCCGCG CAGGACCGCG 980
 CTCCTTTGAG GGTGCGCTCT TGCGACGCG CCACTCCAC CCACTCGTGG CGCGCAGCA CACGCGGGCG 1050
 CCGCCATCCA CATCGCGGCG ACCAGCTCCC TGGGACACG CTGTGCCCCG GGTGTACGCG GAGACCAAGC 1120
 ACTTCTCTTA CTGCTCAGCG GACAAGGAG AGCTCGGGCG CTGCTTCTTA CTCAGCTCTC TGAGGCCAG 1190
 CTGACTGGC GCTGCGAGCG TCGTGAGAC CATCTTTCTG GGTTCAGCG CCTGGATGCG AGGAGCTCCC 1260
 CGCAGGTGCG CCGCGCTGCC CCGCGGCTAC TGGCAATGCG GGCGCTGTT TGGGAGCTG TTGGGAGCG 1330
 ACAGCGAGT CCGCTTACGG GTGCTCTCA AGACGCACTG CCGGCTGCGA GTGCGGCTCA CCGGACGAG 1400
 CGGCTGCTGT CCGCGGAGAG CTGCGTACGG AGCCCGTGGC AGGTGTACGG CTTCGTGGCG GCGTGGCTCG 1470
 CGCGCTGTGT GCGCGGAGCG CTCTGGGCTG CAGGACACAA CGAACGCGAG TTCTCTGAG ACCTACAGAA 1540
 GTTCTATCTT CTGGGGAAGC ATCGCAAGCT CTGCTGTGAG GAGCTGACGT GGAAGTAGAG CGTGGGGAG 1610
 TGGCTTGTG CTGGGGAAGC CCGAGGGGTT GCGTGTGTT CGGCGGAGCA GCACCTGCTG CTCTGAGGAG 1680
 TCGCTGGCA TGTCTGTGAC TGGCTGATGA GTGTGTACGT CGTGGAGCTG CTCAGCTCTT TCTTTTATGT 1750
 CACGAGAGC ACCTTTTCAA GAACAGGCT CATCTGCGGG AGCTGTGCGA AGCAGAGGCT AGGACAGCAT 1820
 ATTGGAATCA GACAGCACTT GAAGAGGCT GACTGCTCAA GACTCCGCTT CATCCCGAAG CGCGCGGCGT 1890
 GGGAGAGCT CAGCGCGCTG CCGCGCTTCA GAGCTGCAAC TACGAGCGGG CGCGCGGCGG GCGGCTCTCT 1960
 TGGGAAGCT GACTAGCTCG TGGAGGCCAG AACGTTCCCG AGAGAAAGA GGGCGGAGCG TCTCACTCTG 2030
 AGGGTGAAG CACTGTTCAG CGGCTCAAAC TACGAGCGGG CGCGCGGCGG GCGGCTCTCT GGGCGCTCTG 2100
 TCGTGGCGCT GGCAGATATC CACAGGGCTT GCGGCACTTT CGTCTGCTG CTGCGGGCGG AGGACCCGCG 2170
 CCGTGAAGCT TACTTTGTCA AGGTGGATGT GACGGGCGCG TACGACACCA TGCCCCAGGA CAGGCTCAGC 2240
 GAGGCTATCG CACAGATCAT CAAGCCAGG AACAGCTACT CGTGTGCTG GTATGCCGTG GTCCAGAAGG 2310
 CGCGCATGT GCGACTGCGG AAGCGTTTCA AGAGCCACGT CTCTACCTTG ACAGACCTCC AGCGGTATAC 2380
 GCGCAGATG GTGCGCTACC TGGAGGAGAC CAGCGCGCTG AGGATGCGG TCGTCATCGA CGAGAGCTCC 2450
 TCGCGAAGT GCGGCGAGCG GAGCTCTTTC GAGCTCTTCC TACGCTTCAT GTGCGCACCT GTGCGCACCG 2520
 TCAGGGGCA GGTCTACGCT CAGTGGCAGG GGATCCCGCA GGGCTCCAT CTCTCCACG CTGCTCGAG 2590
 CCGTGTGCTA GCGGACATGG AGAACAGCT GTTTGCGGGG ATTGCGCGGG ACGGGCTGCT CCGTGGTTTG 2660
 GTGGATGATT TCTGTGTGCT GACACCTCAC CTCACCCAGC CGAAAACTT CTTCCAGGAC CTGGTCCGAG 2730
 GTGTCCCTGA GTATGGTGGT GTGTGTAAGT TCGCGAAGAC GTGCTGTAAC CTTCCCTGTAG AAGACGAGAG 2800
 CTTGGGTGCG ACGGCTTTTG TTAGATGCG GCGCCACGCG CTAATTCCCT GTGCGGCGCT CCGTGTGGAT 2870
 ACCCGAGCCC TGAAGGTGAG GAGCGACTAC TCCAGCTATG CCGCGACCTC CACTAGAGCC AGTCTCACTT 2940
 TCAAGCCGCG TCTCAAGCT GGGAGGAACA TCGCTGCGAA ACTCTTTGGG GTCTTGGCGG TGAAGTCTCA 3010
 CAGGCTGTTT CTTGATTTGC AGGTGAACAG CCTCCAGAGC GTGTGACCA ACATCTACAA GATCTCTGCG 3080
 CTCTAGGCGT ACAGGTTTCA CCGATGTGTG GTGACGCTCC CATTTTCACT GCAAGTTTGG AAGAAGCCCA 3150
 CATTTTCTCT CGCGGCTATC TCTGACAGCG CCGCGCTCTG CTACTCCATC CTGAAAGCGA CAGACAGGAG 3220
 GATGTGCTGT GCGGCGCAGG GCGCGCGCGG CCGCTGCGCC TCGGAGGCGG TGGAGTGGCT GTGCCACCAA 3290
 GCAATCTGCG TGAAGACCGT GTCACCGGT GTCACCTACG TGCCACTCTG GTGCGGCTGA AGCCAGGCG 3360
 AGACGCACT GAGTGGGAG CTCCCGGGGA CGAGCTGAC TGCCCTGGAG GCGCGACGCA ACCCGGCACT 3430
 GCGCTCACAG TCTGGAACCA TCTGGAAGT ATGCGCACCG GCGCACAGCG AGCCGAGAG CAGACAGGAG 3500
 CAGCGCTGTG ACGCGGGCTG CTAGCTCCCA GGGAGGGAGG GCGCGGCGCAC ACCCAGGCGG GCACCGCTGG 3570
 GAGTCTGAGG CTGCTGTGAG GTTTTGGCG AGGCTGCAAT GTCCGGCTGA AGCTGAGTC TCGCGGTGAG 3640
 GCGTGAAGCA GTGTGCGCG CAGCGCTGAG TGTCCAGCAC ACCTGCGCTG TTACTTCCCG CACAGGCTGG 3710
 CGCTGGGCTG CACCCGAGCG AAGGCTTTTC CTCACAGGA GCGCGGCTCT CACTCCCCAC AGTCAAGTAT 3780
 TCCATCCCA GATGTGCCAT TGTTCACCCG CGCGCTGCC CTCTTTGGC TTCCACCCCG ACCATCCAGG 3850
 TGGAGGCCCT TGAAGGAGCC CTGGGAGCTG TGGGAATTTG GAGTGAACAA ATGCTGTGCC TTATACAGAG 3920
 CGAGGACCTT GCACCTGGAT GGGGTGCTCT GTGGTCAAA TTGGGGGGAG GTGCTGTGGG AGTAAATAC 3990
 TGAATATATG AGTTTTTCAG TTTTGA AAAA AAAA AAAA AAAA AA 4042

09582246.092100

Fig. 7

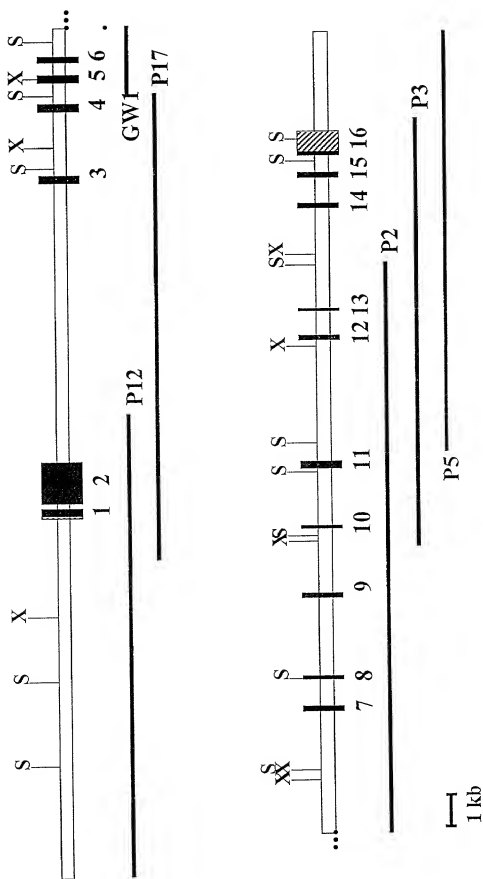


Fig. 8A

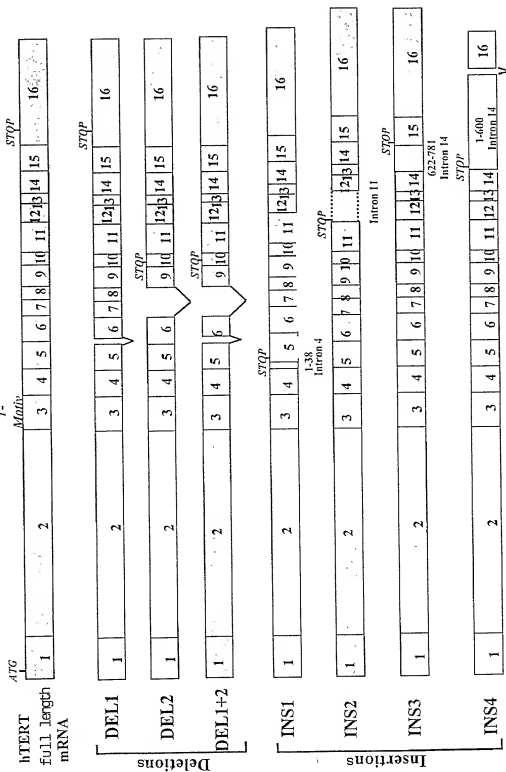


Fig. 8B

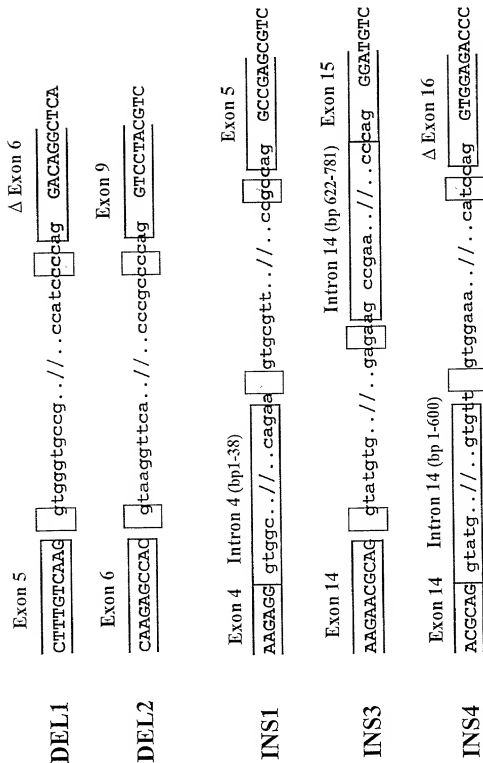


Fig. 9

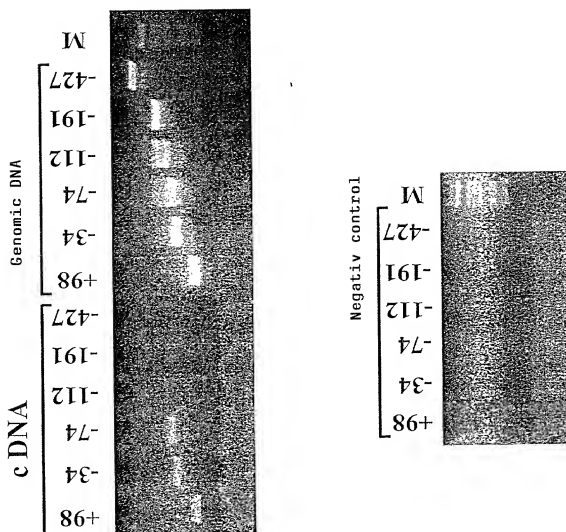


Fig. 10

ACTTGGAGCC	AAGAGTTCAA	GGCTACGGTG	AGCCATGATT	GCAACACCAC	ACGCCAGCCT	TGGTGACAGA	-11204
ATGAGAGCCT	GTCTCAAAAA	AAAAAATAAA	AATTGAAATA	ATATAAAGCA	TCTTCTCGG	CCACAGTGGA	-11134
ACAAAGCCAG	AAATCAACCA	CAAGAGGAAT	TTTGAAATAT	ATACAAACAC	ATGAAATATA	AAACATATAC	-11064
TTCTGAATGA	CAAGTGAAGT	AATTGAAGAA	TTAAAAAGGA	AATTGAAAAA	TTTATTAAAG	CAAAATGATA	-10994
CGGAAACATA	ACCTCTCAAA	ACCCACGGTA	TACAGCAAAA	CGAGTGTCTA	GAAAGGAAGT	TATAGCTATA	-10924
AGCAGGTACA	TCAAAAAAAT	AGAAAAGCCA	GGCCGAGTGG	CTCATCGCTG	TAATCCGAGC	ACTTTGGGAG	-10854
CCCAAGGGGG	GCAGATCGCC	TGAGTCTCAG	AGTTCGAGAC	CAGCCTGACC	AACACACAGA	AACCTTTGCG	-10784
GCTCAAAATA	TACAAATATA	CTGTGGCATG	GTGGCATCAT	CCTGTAATCC	CAGCTCATCT	CGAGGGCTGAG	-10714
GCAGGATACA	CGCTTGAACC	CAGGAGGTGG	AGGTTGCGGT	GAGCCGGGAT	TGCGCCATTG	GNCTCCAGCG	-10644
TGGGTAAACA	GAGTGAACAC	CTGTCTCAAG	AAAAAATAAA	AAATGAGAAA	AGTATAGAAA	AAAGTCTTAAT	-10574
GATGCACCTT	AAAGAACTAG	AAAAAGCAAG	GCACACTAAA	CCTAAATTTG	TGAAAAGAAA	AGAAATATAA	-10504
AAGATCAGAG	CAGAAATATA	TGAAACTGAA	AGATAACAAT	ACAAAAGATC	AACAAAATTA	AAAGTTGGTT	-10434
TTTTGAAATG	ATAAACAAAA	TTGACAAACC	TTTGCCCGAG	CTAAGAAAAA	AGGAAAGAA	ACCTAAATTA	-10364
ATAAAGTCAG	AGATGAAAAA	AGAGACATTA	CAACTGATAC	CACAGAAATT	CAAGAGGATC	CTAGAGGCTA	-10294
CTATGAGCAA	CTGTACACTA	ATAAATTTAA	AACTTAGAAA	AAAATAGATA	AATTTCTGTA	TGCATCAAC	-10224
CTACCAAGAT	TGAACCATTA	AGAAATCCAA	AGCCCCAAAC	GACCAATATC	AATATGGGTA	TAAAGGCTAT	-10154
AAATAAAGAT	CTCTACAGTA	AGAGAAAGCC	AGGACCCAAT	GGCTTCCCTG	CTGGATTTAT	CCAACTCTAT	-10084
AAAGAAGAA	TGAATCCCAT	CTACTCAAAA	CTATTCTGAA	AAATAGAGGA	AAGAATATCT	CCAACTCTAT	-10014
CTCATCTGGC	CAGTATTACC	CTGATTTCAA	AACACACATA	AAAAACAACA	AAAAACAACA	AAAAACAACA	-9944
CAGAAAGAAA	GAAACTACAA	GGCCAAATAT	CCTGATGAAT	ACTGATACAA	AAATCTCTCA	AAAAACAACA	-9874
AGAAACCCAA	TAAACCAACA	CTCTCGAAGG	ATCATTTTAT	GTGATCAAGT	GGGATTTTAT	CCAGGGATCTG	-9804
CGAGTTGGTT	CAAGCATATG	CAACCAATCA	ATGTGATACA	TCACTCCCAAC	AAAATGAAGT	AAAGGATGGA	-9734
TATGATTTAT	TCACTTTTAT	CAGAAAAGGC	ATTGTATATA	ATTCTGCAAC	CTTCATGATA	AAAAACCTCA	-9664
AAAAACCAAG	TATACAGACA	CAGGCACAGT	GGCTCACAGT	CTGCGATTCC	TGGCATCTGG	AGACTCTGGG	-9594
AGGCCAAGGT	GGGATGATTT	CTTGGGCCCA	GGGATTTTGA	ACTAGCCCTG	GCAACCAACT	GAGACTCTGGT	-9524
CTCAAAAAAA	TTTTTTTAAA	AAATTTAGCA	GGCATGATGG	CATATGCTGT	TGATCCCAAG	TGCTCTGGAG	-9454
CGAGGTGTAG	GAGAAATCACT	TAAAGCCTAG	AGGTCGAGGG	TGCGAGTGAG	CATGAACTAT	TCAGCTGTACT	-9384
CCAGCTCAGA	CAACAAGACA	AGACCCCACT	GAATAAGAG	AAGGAGAGAG	AAGGAGGAGA	AGGAGGGGAG	-9314
AGGAGGAGAG	GAGGAGAGGG	AGGAGGTGGA	GAGAGAGTGG	AAGGGGAAAG	GGGAGGAGAG	GAGGAGGAG	-9244
AAGAAACATA	TTTCAACATA	ATTAAGGCC	TATATGACAG	ACCGAGGTAG	TATTATGAGG	AGGATCTGAA	-9174
AGGCTTTTCT	CTAAGATCTG	GAAATAGACA	AGGCGCCACT	TTCAACCACT	CTGATTCAACA	TAGATCTGAA	-9104
AGGCTCTAGT	AGAGCAATCA	TATAGAGAAA	AGAAATAAAA	GGCATCCAAA	TGAAAAGAGA	AGAGTCTGAA	-9034
TTATCTGTTT	TGCAATGATG	ATGATCTTAT	ATCTGAAAAA	GACCTTAAGAC	ACCACTATAAA	AACTATTTAGA	-8964
GTGAAATTTT	GGTATCAGAG	CAATGTACAA	CAATGTACAA	AAATCAGTAG	TCTTCTATA	TTCCACAGCG	-8894
AAACAACTCT	AAAAAGAAAC	CAAAAAGACA	GCTCAAAATA	AAATTAACAA	GCTAGGATAT	AAACCAAGAG	-8824
TGGAAGATCT	TCTCAATAGA	AACTATATAA	ATGTTGATAA	AAGAATTTGA	AGAGGGCAAC	AAAAAGAAAA	-8754
AGATATTCCA	TGTTCAATAGA	TGGAAGAAAT	AAATACTGTT	AAATGTGCTA	ATCTACCCCA	AGCAATTTAC	-8684
AAATTTCAAT	CAATCCCTAT	TAAATACTA	ATGACGTTCT	TCACAGAAAT	AAGAAACAAA	CTTGAGCAT	-8614
TTTGACAGAA	CCACAAGAGA	CCGGAATAG	CCAAAGCTAT	CCTGACCAAT	CATGCTACTG	GCATATAAAT	-8544
CACATTAOC	GACTTTCAAT	TATACTACAA	AGCTATAGTA	AACTAAATCT	ATCATCTACT	GTGATCTAAC	-8474
AGATGAGACA	TGGACAGAG	GAACAGATA	GAGATCCAG	TAATCTCTTC	ATAAATAGT	CGTGGAGGAA	-8404
TTTTTGACAA	AGGTTGCCAG	ACATACTCTT	GGGAAATAAG	CTCTCACCAT	ATACAAAGG	AACTCAAAAT	-8334
CTGGATATCC	ATATGCAAAA	TAACTACTCT	AGAACTCTGT	CTCTCACCAT	AACACCGGAG	AACTCTCCCA	-8264
GGATGAAAGG	CTTAATATCT	TAACCTCATA	CTTTGCAACT	TAATTAAGAA	CACCAACAGG	AAACACAGAC	-8194
GGACATGGGA	GTGGGCAAGG	ACTCTCTGAG	TAATTCCTGT	CGGACACAGG	CACCAACAGG	AGAGACACAC	-8124
AAATGGGATC	ATATCAAGT	AAAAAGCTTC	GGCCAGCACA	AGGAACAACT	CAGGAATATA	ATATTAAGGA	-8054
CCACAGGAAT	GAGGAATATA	TTTGCAAACT	ATTATCTAA	CATGGAATTA	ATAACACGTA	TATATAAGGA	-7984
CTCTCAAACT	CTCTATAAGA	AAAAACCTTA	ATAAGCTGAT	CAAGGATCTA	ATACCAAGTA	TCTGGGTAGA	-7914
CATTCTCCAA	AAATAGTCAAT	ACAAATGGCA	AACAGGCATC	TGAAATTTGT	TAGCAACCCA	CTGATCATCA	-7844
GAGAAATGCA	AAATCAAACT	ACTATGAGAG	ATCATCTCAT	CCGAGTTAA	ATGGCTTTTA	TTCAAAAGAT	-7774
AGGCAATTAAC	AAATGGCAAT	GAGGATGTGG	ATAAAAGGAA	AGCCTTGGAG	AGATTTGGTG	GGATGTGAAA	-7704
TGCTACCAAC	TATGGAGAAC	AGTTTGAAG	TTCTCCAAAA	AACCTAAAAAT	AAAGCTTACC	TACAGCAATC	-7634
CCATTGCTAG	GTATATACTC	CAAAAAGGG	AATCAGTTGT	TCACAAAGCT	ATCTCACCAT	CCACATTTAT	-7564
TGCAGCACTG	TTCTATAGCA	CCAGAGTTTG	GAGCAACACT	CAGTGTCCAT	CAACAGACGA	ATGGAAGAA	-7494
AAATGTGGTG	GCACATACAC	AAGTGGATAC	TAGCCAGCCA	TAAMAAGAA	TGAGATCTGA	CTGAGTGAAA	-7424
CAGCATGGGG	GGCATCGGTC	AGTATGTTAA	GTGAAATAG	CCAGGACAG	AAGACACAA	TTTTCATGTT	-7354
CTCCCTTACT	TGTGGGAGCA	AAATTAATAA	CAATTGACAT	AGAAATAGAG	GAGATGTGTT	TTTTCATGTT	-7284
GGTGGGGGAC	AGGTTGACTA	GAAATCAACA	TAATTTATGT	TATGTTTATA	ATAATCAATA	AGAGTATGAT	-7214
TGGGTTGTTT	TGATGACAAA	GAAAGGATATA	ATGCTTTGAG	TGTCAGATA	CCGCCATTAT	CTGATGTGTA	-7144
TTTATTACAA	TTGTATGCTC	GTATCAAAAT	ATCTGATGAT	TGCTATAGAT	ATAAACCTTA	CTATATTAAA	-7074
AAATTAATTT	TTAATGGCCA	GGCAGCGTGG	CTCATGTGCT	TATCTCCAGC	ACTTTGGGAG	GCAGGGGCTG	-7004
GTGGATCAAC	TGAGGTCAAG	AGGTTTAAAC	GGTGTGACAT	ACCATGATGA	AACCTGCTCT	CTACTAAAG	-6934
TACAAAAAAT	AGCCAGGGGT	GGTGTGACAT	ACCTGTACTC	CAACACTATC	AGGAGGCTGA	CACAGAGAAA	-6864
TTGCTTGAAC	CTGGGAGGGG	GAGGTTGCAG	TGAGCCAGGA	CTATGCCACT	GCACCTGACG	CTGGGTGACA	-6794
GAGCAGAGCT	CCATCTCAAA	ACAAAAACA	AAAAAAGAA	ATTAAATTTG	TATATTTTAT	TACCGGTATA	-6724
AAATATATCT	TACTATATAT	AGAAATTAAT	AAATTAAGAA	ATTATAAAG	GTAATTAACC	ACTTAATCTA	-6654
AAATTAAGAA	AATGTATGTG	GGGTTTCTAG	CTTCTGAGAA	AGTAAAAAGT	ATGGCCACGA	TGGCAGAAAT	-6584

Fig. 10

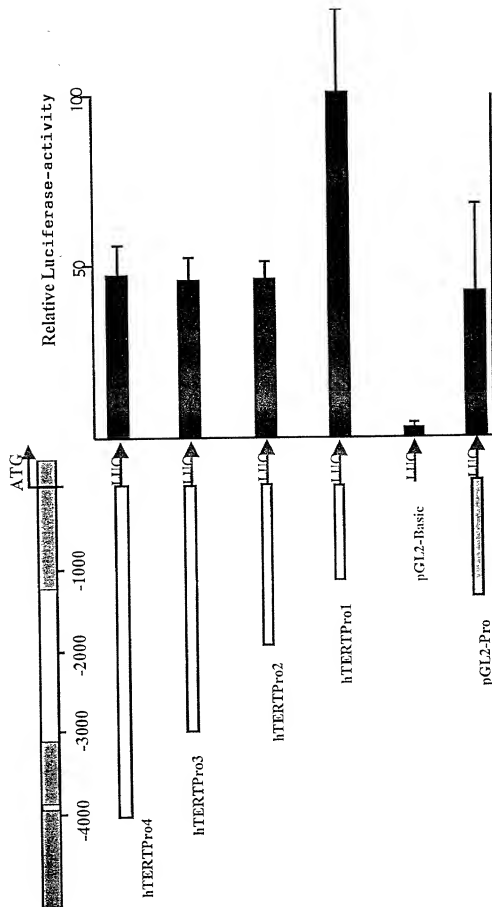
GTGACGAGGG	AACAGTGGAA	GTTACTGTGG	TTAGACGCTC	ATACTCTCTG	TAACTGACTT	AATTTTAAAC	-6514
AARGACAGGG	TGGGAGTAGT	TAAAGAGGCA	TTCTATAAGC	CCTAAAAACA	CTGCTAATAA	TGGTGAAAGG	-6444
TAACTCTCTAT	TAAATTACCAA	TAAATTACAGA	TATCTCTAAA	ATCGAGAGCTG	AGAAATTGGCA	CGTCTGATCA	-6374
CACCGTGTGCT	TGCTTTTCTG	TGCTTTTCTG	CTTGTGTGCT	TGGAGATTTT	CGATTGTGTC	TGCGTGTGTT	-6304
GTTAAACTTA	ATCTGTGTATG	ATCTCGAAGC	GAAGAAATGCT	GGTGATTTCC	TCCAGGAAGA	TTAGAGTACC	-6234
TGCAGGGAAG	CAGGTGGGCTG	TGTGGACCTG	AGGCACTTCA	ATCTTCAAGG	CTCTGTGGCC	AAGACCCAGG	-6164
CGTCAGGCGAG	AGGCGCTGATG	ACCCGAGGAC	AGGAAAGCTC	GGATGGGAAG	GGGACGTAGG	AGGCTCTGCC	-6094
TGCTGGTGTG	CAGCGCATAG	AGGCGCTTTA	TTTACGCTTG	GCAGAGATTG	CTCTGGATAG	TGCTGGAAA	-6024
AGGCGGCCAG	CGGGAATGCA	AGGAGTCAGA	AGGCTCTCTG	TCAAACCCAG	GCCACGAGCT	ATGGCGCCCA	-5954
CGCTGGTGTG	CAGCGCATAG	AGGAGAGTCA	AGGCACTCTG	AAGTATGGCT	TAAATCTT	TTTCACTGGA	-5884
AGGCGGCCAG	CGGGAATGCA	AGGAGTCAGA	AGGCTCTCTG	TCAAACCCAG	GCCACGAGCT	ATGGCGCCCA	-5854
AGCAGCTGACC	AAGGTGTATT	CTGAGGGGAG	CTTGAGTTAG	TGCTCTCTCT	TAAACAGAAA	AGTCATGGAA	-5814
GCACCTTCTT	CAAGGTGAAA	CCGACAGCCC	GCTCTGCGGT	CATTACCTTC	TTTCTCTCT	CCCTCTCTTG	-5744
CGCTCGCGGT	TTCTGATCGG	GACGAGGTGA	CCCGCTGGA	GCTTCTCCGA	CGCCGTGCTG	AGGACCCCTT	-5674
CCCGGGCGTG	TGCCAGAGGG	AGGAGAGTCA	AGGCACTCTG	AGATATGGCT	CTTAATTAACA	AACCTGGGATG	-5604
TGCGTGGGGG	CGGAGACGGA	CGGCGGGATT	CAAAGACTTA	ATTCCATGAG	TAAATTTCAAC	CTTTCACCAT	-5534
CGGAATGGAT	TGATGTTTTA	TCTTAATATT	TTCTTAAATT	TCATCAAAAT	ACATTCAGGA	CTGCAGAAAT	-5464
CCAAATGGCT	AAAACAGGAA	CTGAGCTATG	TTTGCCAAAG	TCCAAGGACT	TAAATTAACAT	TGTCAGAGGG	-5394
OTTTTTCGCT	TAAAGTATT	TTTATTGGTT	TTTATAAGGT	GGCTTGGGTT	GCAGAGGAAA	GTACAGAGGG	-5324
AGAGGCGCTG	GGCGGAGGCG	TATGAGCAGC	GCAGGGCCAC	CGGGGAGAGA	TGCCCCGGCC	TGGGAGGCTG	-5254
ACAGCAGGAG	CAGTGACCTG	CTCTCCCTGG	AGCTGGCCACA	TGGGCAAGC	CGAAGCGGCG	CAGCTCGGTT	-5184
GCTGACTCAGG	ACCCCATACC	GGCTTCTCTG	GCCCAACCCAC	ACTAACCAGC	GAAGTCAGGG	AGCTCTGAAC	-5114
CCGTGGAACAC	GAACATGAGC	TTTGGCTGCC	TGCTTCCCTG	GGTGGGTCAA	GGGTGAATGA	GGTGTGTGCA	-5044
GGAAATGGCC	GATTAATTA	CACGACTCTG	CTGATGGGGA	CGGTCTCTTC	CATCATTTAT	CATCTCTACC	-4974
CCCAAGGACT	AAATGATTCC	AGCAACTTCT	TGCGGTGTGA	CAGGCGATGA	CATCAACCTAG	TACAAACACG	-4904
ACTCTTTTAC	TAGGCCACCA	GAGCAGGSCG	CACACCCCTG	ATATATTAA	AGTCCAGGAG	TAGATGAGCT	-4834
GGTTTCCAGC	ACCGAGCTGG	GGTGACACAA	CGCGCTGAAC	AGTCTGTGTC	TCTAGACTAG	AGAGACCTTG	-4764
CAGGACTCCT	CCAGAGTTCT	AGGCGCTGGT	TGCTGCTTCC	CGAGGGGCGC	ATCTGCCCTG	GAGACTCAGC	-4694
CTGCGGTGCC	ACACTGAGGC	CAGCGCTGTC	TCCACACCTC	CGGCTCCAG	CGCTCAGCTT	CTCCAGCAGC	-4624
TTCTTAACAC	CGGTTGGGCG	CTGTGTCAG	CGCTACTGTC	TCACTGTGTC	CACTGTGTCT	TGCTCTGAGC	-4554
ACGTAGCTCG	CAGGCTTCTCT	CGTTCACATG	GGTGTCTGTC	TCTTCCCAAC	ACACTTCAGT	CGGTTGAGCG	-4484
GAGGAGATTG	TGGCGCTTCC	AGACTGSGCT	CTCTGAGGCT	GAACCTTGCT	CGTGGGCCCT	GATGAGGTT	-4414
CTTGGGCTGC	GGGTGCAAGC	TGACCTCCAT	TCCAGGGGCG	TCCGCTCTCT	CTCTGATCTC	CTCGGCGCTG	-4344
CCGCGTGTCT	CTTCTGTTTC	TGCTGCTCTT	TCCAGCTCCA	GCTGCGTGTG	TCTCTGCTCG	CTGAGGCTCT	-4274
GGGTTTGTTA	TAGGCATAGG	ACGGGGGCGT	GGTGGGCGAG	GGGCGCTCTT	GATGAGGCGA	CATCTGGGTC	-4204
TGAATAGTAG	AGTGCCTGTC	CTCACTTAGG	TCCACGGGCA	CAGGCGTGGG	GATGAGGAGCC	CGCCACAGGA	-4134
CCCGCCCTTC	CTTGCCCAAC	ACTTCTCTCG	CCCGCTCCCT	CTGGAACATT	GAGTGGGAGT	TTCCACAACT	-4064
ACTAGCATC	CTCTTCCCAA	AAGACCCAGC	ATTGGCACCC	CTGGACATTT	GCCCACACAG	CTTGGGAATT	-3994
c-Myc							
CAGAGTGACTA	CGCACAATAT	TAGACCACTC	CGGTCCACGA	CCGACCCCGC	CTGTTTATTT	TTAATAGCTA	-3924
CAAGGACAGGG	AATACCTCTG	TTTAAAGTGC	TTTAAACAAC	TGGTTAAACA	ACGGTGGTCA	TCCGCAAGCT	-3854
GGACAGTCTCC	TACAGATGAA	GAGGAACATG	CCGTTTATAA	AGGCTCGAGG	CATCTCAGAG	GAATTAACGT	-3784
GCTACAAACAG	TGCCACCTCC	TAGGATACAG	TAGGCAACAT	GCTCAAAAGT	ATAGATATTC	ACCCCTATGG	-3714
AGGGGAGTGG	TTAGGGGGGT	TAAAGACGGT	GGGGCGGGCA	GGGTGGGGCT	CTCTGCAAGC	ACCTTTACTA	-3644
AAGCGAGTTT	CTCTGGTCTG	TAGGTATTGG	CTCAGTTATG	GGAGACTAAC	CATAGGGGAG	TGGGGATGGG	-3574
AGCCACCGGGA	GGCTGTGCCA	CTTGTGCCAT	GGCCGAGTGT	CCTGGGCGAG	ATAATGCTCT	AGAGATGCC	-3504
AGTCTCTGAT	TCCCCAAAC	TGCTGGACAG	AACCCCGGCG	GGCCAGGGCC	GTATTGAGT	GTGATCTCCG	-3434
TGAGGACCTT	GAGGTCTGGG	ATCTCTGGG	ACTACTTGCA	GGCCGAGAAA	CTTATCAGG	GGTTCTGGGA	-3364
AGAGGGGGGG	AGAGGGGGCA	GAGGGGGGCA	GCTCAGGAG	GATGGGAGCA	GTCAGTCTGA	GGCTGAAAG	-3294
GGAGGGAGGG	CTCGAGGCC	AGGCTGCAAA	GGGCTCCAG	AAGCTGGAAA	GGAGGGGGA	GGGACCTCC	-3224
ACGGAAGCTG	CAGCAGGAG	CGGCGGCTG	CCCTTAGGCC	ACCAAGGGCC	ATTCGAGAAC	TCCGCGCTCC	-3154
TGTGCATAGG	AGGCGCATCG	CGCTGCCCTT	CTAGCATGAA	GTGTGTGGGG	ATTTGCAAGA	CAGGCGCTGAA	-3084
ACCCATGCAC	TGTGAATCTA	GGATTATTTC	AAAACAAGAG	TTTACAGAAA	TATTTTAT	TCTACTTCTT	-2944
TGTGCTCCGG	GCAAGGGGAC	GCGAGGCACG	AGTGAATTTA	TTAGGACTTA	CTCTGGCTCA	CTGCACCTTC	-2874
CTGAGACAGA	GTTATGCTCT	TGTTGGCCAC	GCTGGAGTGC	AGAGGACTTG	GATTTACAGG	CTGCACCAAC	-2804
CGTCTCTGCG	GTTCAAGCAA	TTTCTGTGCC	TGAGGCTCCC	CACCATGTGG	GATTCAGCTGA	TCTCAAAATC	-2734
ACACCCGGCT	AATTTTGTAT	TTTCTGTAGA	GATGGGCTTT	ACATGCTTGC	AGGCTAGAGC	CATCACTCTC	-2664
CTGAGCTCAG	GTTATCTCGC	CACCTCAGCC	TCCCAAAGTG	CTGGGATATC	GATGAGAGTT	CATGAGGATT	-2594
GGCCTATTTA	ACCATTTTAA	AACCTTCCCT	GGCTCAAGTC	ACACCCAGCT	TTCTAGACT	GAGCATATAC	-2524
AATTTTCCCT	TTAATCAGGA	GTTTACCTCT	TTTGATATTT	TCTGTAAATC	CCCATCTGAG	GGGACGCTG	-2454
CGTCTCTTGA	CATATTACCA	GTATCTGTGA	CCACCTGTTA	TCCGATGGA	GTAGTATGAG	ATCAGGCGGC	-2384
GAGGCTGACG	GCTTCAGGTC	CCAGTGGGGT	TGCCATCTGC	CAGTGAACAT	GATGAATTTA	AAGTCCATCT	-2314
AAGTGTGGAC	ACTGTCCTGA	ATGCTCAATG	CTCAGTGTGC	CTCGGAGAC	GGATTTTCT	CTCATCTCTG	-2244
CTCTCACTCT	ACTGGGATAT	AGGCCCTTTC	CTATGCCCTCC	CCAGGGGACG	CTGTTTCAAT	TGTTGGTGTG	-2174
TGGAGGAAGG	AATGATACT	TGTTATATTT	CAGTGCCTTC	ACTGATATCA	CGCTGGAGGG	AGTCAATAGT	-2104
TTTGTTTTGT	TTTGAAGAGG	GTTTTCACCT	TTTGTGCTCA	GGCTTCCGCG	CTCCCATTTG	GCTGGGATTA	-2034
GCTTACTGCA	GCTCTGCTCT	CCCAGGTTCA	AGTGAATTCT	CTGGACATTT	GGGGTGGGTT	GGGTTTCAAC	-1964
CAGGACCCCG	CCACATGACC	CAGCTAATTT	TTTGTATTTT	TAGTAGAGAC			

- 14/15 -

Fig. 10

ATGTTGGCCA GGCTGGTCTC GAACCTCTGA CCTCAGATGA TCCACCTGCC TCTGCCTCCT AAAGTGCTGG -1894
 GATTACAGGT GTGAGCCACC ATGCCACGCT CAGAATTAC TCTGTTTGA AACATCTGGG TCTGAGGTAG -1824
 GAAGCTCACC CCACTCAAGT GTTGIGGTGT TTTATGCCAA TATAGAATT TTTTATTGT TGTAGAACA -1754
 CTCTTGATGT TTTACACTGT GATGACTAAG ACATCATCAG CTTTTCAAAG ACACACTAAC TGCACCCATA -1684
 ATACTGGGGT GTCTTCTGGG TATCAGCAAT CTTCATTGAA TGCCGGGAGG CGTTTCTCTG CCAATGCACAT -1614
 GGTGTTAATT ACTCCAGCAT AATCTTCTGC TTCCATTCT TCTCTTCCCT CTTTAAAAAT TGTGTTTCT -1544
 ATGTTGGCTT CTCTGCAGAG AACCAGTGTA AGCTACAAC TAACTTTGT TGGAAACAAT TTTCAAACCC -1474
 GCCCTTTTGC CCTAGTGGCA GAGACAATC ACAAACACAG CCCTTTAAAA AGGCTTAGGG ATCACTAAGG -1404
 GGAATTCTAG AAGAGCGACC TGTAACTCTA AGTATTTACA AGACGAGGCT AACCTCCAGC GAGCGTGACA -1334
 GCCCAGGGAG GGTGCGAGGC CTGTTCAAAT GCTAGCTCCA TAAATAAAGC AATTTCTCTC GGCAGTTTCT -1264
 GAAAGTAGGA AAGGTTACAT TTAAGGTTGC GTTTGTTAGC ATTTCAAGTGT TTGCCGACCT CAGCTACAGC -1194
 ATCCCTGCAA GGCCTCGGGA GACCAGAAG TTTCTCGCCC CCTTAGATCC AAACCTGAGC AACCCGAGT -1124
 CTGATTCTCT GGGAAGTCTT CAGCTGTCTT CGGGTTGTGC CGGGGCCCCA GGTCTGAGG GGCACCTGG -1054
 CGGTGTGGCT TCTACTGCTG GGCTGGAAGT CGGGCTCTCT AGCTCTGCAG TCCGAGGCTT GGAGCCAGGT -984
 GCCTGGACCC CGAGGCTGCC CTCACCCCTG TCGGGCGGGG ATGTGACCAG ATGTTGGCCT CATCTGCCAG -914
 ACAGAGTGCC GGGGGCCAGG GTCAAGGCCG TTGTGGCTGG TGTGAGCGCG CGGTGCGCG GCCAGCAGGA -844
 GCGCCTGGCT CCATTGCCA CCCCTTCTCG ACGGAGCGC CGGCTGGGT GATTAACAGA TTTGGGGTGG -774
 TTTGCTCATG GTGGGGACCC CTCGCCGCTT GAGAACCTGC AAAGAGAAAT GACGGGCTGT GTCTAAGGAG -704
 CCCAAGTCGC GGGGAAGTGT TGCAGGGAGG CACTCCGGGA GGTCCCGCTT GCCCGTCCAG GGAGCAATGC -634
 GTCTCTGGGT TCGTCCCGAG CCCGCTCTAC GCGCCTCCGT CCTCCCTTTC ACGTCCGGCA TTCTGTTGTC -564
 CCGGAGCCCG ACGGCCCGCG TCCGACCTG GAGGCAGCCC TGGGTCTCCG GATCAGGCCA GCGGCCAAG -494
 GGTGCGCGCA CGCACTGTT CCCAGGGCTT CCACATCATG GCCCTCCCTT CGGGTTACCC CACAGCCTAG -424
 GCCGATTCCA CCTCTCTCCG CTGSGGCCCT CGCTGGCGTC CTTGCACCTT GGGAGCGCA GCGCGCGCG -354
 GCGCGGGAAG CGCGGCCAG ACCCCCGGTT CGGCGCGGAG CAGCTGCGCT GTCCGGGCCA GCGCGGCTC -284
 CCAGTGGATT CGCGGCCACA GAGCGCCAGG ACCGCGCTCC CACGCTGCG CAGGGAGTGG GGACCGGGC -214
 ACCGCTCTG CCCCTTCAAC TTCCAGCTCC GCCTCTCCG CGCGGACCC GCGCGTCC GACCCCTCCC -144
 GGGTCCCCG CCCAGCCCC TCCGGGCCCT CCCAGCCCT CCCCTTCCCT TCCGCGGCCG GCGCGCTCC -74
 TCGCGGCGCG AGTTTCAAGC AGCGCTGCGT CTTGCTGCG CAGTGGGAAG CCCTGGCCCC GGCACCCCC -4
 GCGATG

Fig.: 11



As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name. I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

REGULATORY DNA SEQUENCES OF THE HUMAN CATALYTIC TELOMERASE SUB-UNIT GENE, DIAGNOSTIC AND THERAPEUTIC USE THEREOF

the specification of which is attached hereto,

or was filed on **December 22, 1998**

as a PCT Application Serial No. **PCT/EP98/08216**

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims.

I acknowledge the duty to disclose information which is material to the patentability of this application in accordance with Title 37, Code of Federal Regulations, §1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

Prior Foreign Application(s), the priority(ies) of which is/are to be claimed:

197 57 984.1
(Number)

Germany
(Country)

December 24, 1997
(Month/Day/Year Filed)

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose the material information as defined in Title 37, Code of Federal Regulations, §1.56 which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

(Application Serial No.)

(Filing Date)

(Status)

(patented, pending, abandoned)

(Application Serial No.)

(Filing Date)

(Status)

(patented, pending, abandoned)

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

5- Jeffrey M. Greenman, Reg. No. 26552
 Barbara A. Shimei, Reg. No. 29862
 William F. Gray, Reg. No. 31018
 Alice A. Brewer, Reg. No. 32888
 Jerrie L. Chiu, Reg. No. 41670

all of Bayer Corporation, 400 Morgan Lane, West Haven, Connecticut 06516

Send Correspondence To:
Mr. Jeffrey M. Greenman
Bayer Corporation
400 Morgan Lane
West Haven, Connecticut 06516

Direct Telephone Calls To:
 (203)812-3964(Jerrie L. Chiu)

FULL NAME OF SOLE OR FIRST INVENTOR <u>Gustav Hagen</u>		INVENTOR'S SIGNATURE <i>Gustav Hagen</i>	DATE <u>16 Mai 2000</u>
RESIDENCE D 51373, <u>Leverkusen</u> , Germany <u>DEX</u>		CITIZENSHIP German	
POST OFFICE ADDRESS c/o Bayer Aktiengesellschaft, D 51368 Leverkusen, Germany			
FULL NAME OF SECOND INVENTOR <u>Maresa Wick</u>		INVENTOR'S SIGNATURE <i>Maresa Wick</i>	DATE <u>16 Mai 2000</u>
RESIDENCE D 51065 <u>Köln</u> , Germany <u>DEX</u>		CITIZENSHIP German	
POST OFFICE ADDRESS c/o Bayer Aktiengesellschaft, D 51368 Leverkusen, Germany			
FULL NAME OF THIRD INVENTOR <u>Dmitry Zubov</u>		INVENTOR'S SIGNATURE <i>Dmitry Zubov</i>	DATE <u>16 Mai 2000</u>
RESIDENCE D 51061 <u>Köln</u> , Germany <u>DEX</u>		CITIZENSHIP Russian	
POST OFFICE ADDRESS c/o Bayer Aktiengesellschaft, D 51368 Leverkusen, Germany			
FULL NAME OF FOURTH INVENTOR		INVENTOR'S SIGNATURE	DATE
RESIDENCE		CITIZENSHIP	
POST OFFICE ADDRESS			
FULL NAME OF FIFTH INVENTOR		INVENTOR'S SIGNATURE	DATE
RESIDENCE		CITIZENSHIP	
POST OFFICE ADDRESS			
FULL NAME OF SIXTH INVENTOR		INVENTOR'S SIGNATURE	DATE
RESIDENCE		CITIZENSHIP	
POST OFFICE ADDRESS			
FULL NAME OF SEVENTH INVENTOR		INVENTOR'S SIGNATURE	DATE
RESIDENCE		CITIZENSHIP	
POST OFFICE ADDRESS			

Le A 32 805-US